مقدمة فى

الإحصاء الاجتماعى

تأليف

أ. د. اعتماد محمد علام
أستاذ علم الاجتماع كلية البنات - جامعة عين شمس



مكتبـــة الأنجلـــو المصريـــة

مقدمة الطبعة الثانية

يسعدنى أن أقدم هذا المؤلف فى الإحصاء الوصفى (مبادئ الإحصاء) لأبنائى الطلـبة والطالـبات الدارسين فى مجال علم الاجتماع. وقد راعيت فى تبويب هذا المؤلـف أن يغطـى فى بساطة الأساليب الإحصائية الوصفية، وأن يقدم نماذج من تطبـيقات كـل منها بما ييسر على الطلاب فهم استخدام الإحصاء فى الكشف عن العلاقات بين الظواهر، وتفسيرها بعد قياسها، مع اكساب الطالب القدرة على التنبؤ فـى معالجته الإحصائية للظاهرة الاجتماعية منفردة أو فى علاقتها بظاهرة أخرى. وهـو أمر على جانب كبير من الأهمية لأن التنبؤ هو غاية كل علم كما نعلم. وتبدأ منهجـية العـرض فـى هذا المُؤَلَّف بشرح لأهمية استخدام الإحصاء فى البحوث الاجتماعـية. إذ يبدأ الباحث بجمع البيانات حول ظاهرة ما فى المجتمع ثم يقوم من خـلال الأسـاليب الإحصـائية الملائمـة بتبويـب هذه البيانات وتحليلها إحصائيًا لاستخلاص عدد من المؤشرات.

وتتمـثل أهمية هذه الأساليب الإحصائية فى البحوث الاجتماعية فى أنها تمكن الدارس مثلاً من أن يتعرف على خصائص مجتمعه المحلى من حيث: نسبة الأمية، نسـبة الحاصلين على مؤهلات عليا وتخصصاتهم العلمية المختلفة، متوسط الدخل علـى مستوى المجتمع المحلى، أو على المستوى القومى... الخ. أيضًا نستطيع من خـلال استخدام الأساليب الإحصائية أن نستخلص من البيانات عددًا من المؤشرات الاجتماعـية الـتى تفيد صُنَّاع القرار فى مجال التخطيط والتنمية فى شتى مناحى الحياة.

أيضًـا يعتبر هذا المؤلف الطبعة الثانية لمؤلفنا الموسوم "مقدمة فى الإحصاء الاجـتماعى". وتتمـيز هذه الطبعة عن سابقتها بالعديد من التعديلات، التى تشمل جوانـب كثـيرة من التنقيح والتبسيط والتحديث فى ضوء الخبرة التدريسية الطويلة لمـادة الإحصـاء لطلـبة قسـم الاجتماع وممن يلتحقون بالدراسات العليا فى هذا التخصص. ويضم هذا المؤلف الموضوعات الآتية :

١- المفاهيم الأساسية فى مجال الإحصاء.
٢- تبويب البيانات.
٣- العرض البيانى للبيانات.
٤- مقاييس النزعة المركزية.

٥- التشتت والالتواء.
٦- منحنى التوزيع الاعتدالى والمعايير والالتواء.
٧- الارتباط.
٨- الانحدار.

والله الموفق

أ.د. اعتماد محمد علام

القاهرة فى سبتمبر ٢٠٠٩

أهداف المؤلف :

ويهدف هذا المؤلف إلى أن يعرف الطالب :

١- نـــوع البـــيانات (كمية أم كيفية) ومصادرها وأنواع المتغيرات الكمية (متصلة ومتقطعة). أيضًا كيفية جدولة البيانات الخام وكيفية عرضها بيانياً بعد معرفته بالأشكال البيانية الملائمة للمتغيرات المتصلة والمتغيرات المتقطعة.

٢- مستويات القياس (الاسمى، الرتبى، الفاصلة والنسبة) وخصائص كل منها وأن يستطيع الطالب المقارنة بينها.

٣- مقاييس النزعة المركزية (المتوسط الحسابى، والوسيط، والمنوال) والمعادلات الرياضـــية الـــتى تســـتخدم فـــى حساب كل مقياس منها على حدة، سواء من البـــيانات الخام أم الجداول التكرارية. كما يتم تدريب الطالب من خلال الأمثلة الـــتى تستخدم هذه المعادلات فى الحل وإيجاد المطلوب. أيضًا يتدرب الطالب على إيجاد قيمة كل من الوسيط والمنوال باستخدام الرسم.

٤- مقايـــيس التشـــتت لكل من المتغيرات المتصلة (المدى، والانحراف المتوسط، والانحراف الربيعى، والتباين والانحراف المعيارى ومعامل التباين) ومقاييس التشـــتت للمتغيرات المتقطعة (نسبة التباين ودليل التباين الكيفى). ومن خلال الأمثلة والتطبيقات على كل مقياس على حدة بحيث يستطيع الطالب أن يقارن بينها لمعرفة مزايا وعيوب كل مقياس. وكيفية استخدامه فى وصف خصائص مجتمعه المحلى.

٥- المنحنى الاعتدالى وخصائصه والمقصود بكل من الدرجة المعيارية والدرجة التائية واستخداماتهما.

٦- الالـــتواء وكيفية حسابه بعد معرفة ما درسه الطالب لمقاييس النزعة المركزية ومقايـــيس التشتت. وأن يستطيع الطالب أن يرسم شكل التوزيع وتحديد اتجاه الالتواء حسابياً سواء كان الالتواء ناحية اليمين أم ناحية اليسار.

٧- العلاقـــة بين متغيرين وكيفية رسم العلاقة من خلال الشكل الانتشارى وحساب حجم العلاقة وتحديد اتجاهها باستخدام إحدى معاملات الارتباط الملائمة والتنبؤ بقيمة (ص) عندما تكون قيمة (س) معلومة وباستخدام معادلة الانحدار.

هـــذا ويشتمل كل فصل على أمثلة محلولة وبعض التطبيقات. وقد روعى فى نهايـــة معظم فصول هذا المؤلف أن تشتمل على تعريف مختصر للمفاهيم الأساسية الـــتى وردت فى متن هذه الفصول. كما روعى فى خاتمة كل فصل أن تضم عددا مـــن التماريـــن وتغطـــى فى الوقت ذاته جميع الموضوعات التى يشتمل عليها كل فصل من فصول هذا المؤلف.

١

الفصل الأول

المفاهيم الأساسية ومستويات القياس

مقدمة

١- تعريف الإحصاء.

٢- الأساليب الإحصائية.

٣- تعريف البيانات ومصادرها.

٤- أنواع المتغيرات.

٥- المجتمع الأصلي والعينة.

٦- مراحل البحث الإحصائي.

٨- خصائص التوزيع التكراري.

الفصل الأول
المفاهيم الأساسية ومستويات القياس

مقدمة :

إذا تـــأمل الإنســـان ما يدور حوله من أحداث وتغيرات ومعلومات مقروءة أو مرئـــية أو مسموعة، فسوف يجد نفسه محاطًا بالبيانات الإحصائية. بل إنه يجد مثل هـــذه البـــيانات الإحصـــائية متضـــمنة فى خطاب الحياة اليومية فنظرة واحدة فى الصـــحف اليومية نجد أنها تطالعنا ببيانات إحصائية فى شكل جداول أو رسومات بيانـــية أو نسب مئوية حول البطالة فى سوق العمل، أو تصاعد أسهم فى بورصة الأوراق المالية، أو نسبة الحوادث ومعدلاتها على الطرق خلال عام أو خلال فترة زمنية معينة.

ومن ثم نقول إن الإحصاءات أصبحت جزءًا هامًا من حياة الإنسان، إلا أنها قد تشـــير إلى موضوعات مختلفة من خلال رؤية وأساليب متباينة بين فرد وآخر، وبين باحـــث فـــى مجـــال علمى ما أو باحث فى مجال علمى آخر. فمثلاً، يناقش خبراء الطقس الإحصائيات اليومية حول ارتفاع درجة حرارة الجو أو انخفاضها واحتمالات ســـقوط الأمطار ونسبة كثافتها خلال الأيام القادمة، وسرعة الرياح، والنسب المئوية لرطوبة الجو، وحالة البحر من مد وجذر كل ذلك فى شكل إحصائيات وصفية لما تم رصـــده بـــالفعل عن طريق أجهزة الرصد والقياس، واستخدام الاحتمالات فى توقع الأحوال الجوية المستقبلية حتى باستخدام الأقمار الصناعية التى تعتمد اعتمادًا أساسيًا علـــى البيانات الإحصائية، وينطبق هذا القول على خبراء الرياضة حيث يستخدمون النسب والبيانات الإحصائية فى الوصف والتعليق على مباريات كرة القدم.

ومـــن جهـــة أخـــرى، يختلف أسلوب الخطاب الإحصائى للباحثين فى العلوم الإنســـانية عنه فى العلوم الرياضية. فالباحث فى مجال العلوم الإنسانية يبحث عن الأدوات الإحصـــائية الملائمـــة لتحليل البيانات التى يجمعها حول ظاهرة معينة أو متغـــير ما. أما المشتغلون بالعلوم الرياضية فإنهم يصفون الإحصاء كجزء أساسى من علوم الرياضيات.

وأمـــا من الناحية التطبيقية فنجد أن مجال الإحصاء يعم مختلف التخصصات العلمـــية رغـــم التبايـــن فيما بينها من طب، وصحة عامة، وإدارة الأعمال، وعلم الاجتماع وعلم النفس... إلخ.

يهـدف هـذا الفصل إلى إفهام الدارس الفرق بين مفهومى الإحصاء الوصفى والإحصـاء الاستدلالى وأن يلم بمشتملات الإحصاء الوصفى تفصيلاً - موضوع هـذا المؤلف -. كذلك يهدف إلى تعريف الدارس بالمفهومات الأساسية : البيانات الخـام، البـيانات الكيفية، والبيانات الكمية وأنواعها: متصلة ومتقطعة، والمجتمع الأصـلى والعينة ومستويات القياس المختلفة وخصائص كل مستوى على حدة وأن يقارن بين هذه المستويات وتحديد الأساليب الإحصائية الملائمة لكل مستوى، فضلاً عن التوزيع التكرارى وخصائصه ومعنى الالتواء.

أولاً: تعريف الإحصاء :

تُعَرِّف الإحصاء بأنها مجموعة من الأساليب التى تستخدم فى تجميع ووصف وتحلـيل البيانات الرقمية الدالة على جوانب متعددة ومتباينة للواقع الاجتماعى كما تُعَرَّف الإحصاء بأنها تشكيلة من النظرية والمناهج يتم تطبيقها بغرض فهم البيانات وتقديم دلالة ميدانية على قبول أو رفض الفروض المشتقة من النظريات المستخدمة فى العلوم السلوكية.

ثانيًا : الأساليب الإحصائية :

تنقسم الأساليب الإحصائية إلى قسمين أساسيين هما :

١- **الإحصـاء الوصـفى Descriptive statistics**: ويـتألف مـن مجموعة الأساليب التى تصف الظواهر الاجتماعية من خلال أوصاف رقمية. فمثلاً إذا أردت أن تصـف مجـتمعك المحلـى الذى تعيش بداخله بدلالة ثلاثة متغـيرات هى الجنس Sex، العمر Age، ودخل الأسرة، فإن الغرض من هـذا الوصـف هو تقديم وتفهم جيد لعدد كل من الذكور والإناث، والبنية العمـرية للمجـتمع المحلـى، ونسـب الأسر التى تقع داخل فئات الدخل المخـتلفة. ولكـل متغـير خواص محددة. فعلى سبيل المثال يشير توزيع الأفراد المقيمين داخل هذا المجتمع المحلى وفقًا لمتغير الجنس إلى عدد كل من الذكور والإناث.

٢- **الإحصـاء الاستدلالى Inferential statistics**: يتألف من مجموعة الأساليب الإحصـائية الـتى يستخدمها الباحثون فى الاستدلال على خصائص المجتمع الأصـلى مـن خـلال المشـاهدات التى يتم إجراؤها على عينة ممثلة لهذا المجـتمع. فمثلاً، يُستخدم الإحصاء الاستدلالى فى وصف المجتمع الكبير من

خـــلال اســـتخدام الباحثين للمعلومات من عينات صغيرة الحجم نسبيًا من هذا المجـــتمع. وتنهض معظم البحوث الاجتماعية على الإحصاء الاستدلالى نظرًا لصعوبة دراسة المجتمع الأصلى، وتكلفة البحث الباهظة ماليًا، وما قد تتطلبه من جهد شاق.

فلـــو افترضـــنا أنك أردت دراسة مجتمعك الذى تعيش فيه ويبلغ عدد أفراده ٤٠٠٠٠٠ نســـمة، فهل فى مقدورك أن تسأل كل فرد من أفراد هذا المجتمع حول سنه ودخل أسرته والتعرف على نوعه؟ وكم تحتاج من الوقت والجهد الشاق للقيام بهـــذا البحـــث؟ فضـــلاً عـــن التكلفة المالية التى يتم إنفاقها على فريق الباحثين أو المعاونيـــن لـــك؟! لذلـــك من الأفضل أن يلجأ الباحث إلى اختيار عينة ممثلة لهذا المجـــتمع ولنفترض مثلاُ أنها تضم ٢٥٠ فردًا. ففى هذه الحالة يسهل على الباحث أن يجمـــع من أفراد العينة جميع المعلومات الخاصة بالسن والدخل والنوع. ويمكن مـــن خلال هذه المعلومات الاستدلال على شكل التوزيعات للمتغيرات الثلاثة على مســـتوى المجتمع الأصلى. ولو كان اختيار العينة صحيحًا والوسائل المستخدمة فى جمـــع البيانات ملائمة لطبيعة المجتمع وأهداف البحث، فسوف يحصل الباحث على وصف دقيق لخصائص هذا المجتمع.

ولما كان الوصف الرقمى للعينات والمجتمع الأصلى ذا هدف متماثل، كان ضروريًا على الباحث أن يحافظ على ما يميز بينهما عندما يشير إلى أى منهما. فمثلاً من خلال الاستدلال على خصائص المجتمع الأصلى نقول إن القيم تكون ثابتة بمعنى أن عدد كل من الذكور والإناث داخل المجتمع الأصلى فى أى وقت وعـــند أى نقطة ستمثل قيمة ثابتة لا تتغير ولتكن مثلاً ٣٩% للذكور مقابل ٦١ % للإنـــاث. هـــذا الوصف الرقمى للذكور أو الإناث داخل المجتمع الأصلى، يعـــرف احصـــائيًا بـــالمَعْلم The parameter (Kurtz, 1983: 2) وتبدو أهمية الإحصـــاء فى وصف العينة الذى لا يعطى قيمًا ثابتة إذا قام الباحث بأخذ أكثر مـــن عيـــنة مـــن هذا المجتمع فسوف يحصل الباحث على نسب مئوية مختلفة للذكـــور مـــن عيـــنة لأخرى. فمثلاً قد تعطى خصائص العينة الأولى أن نسبة الذكور ٣٨,٢%، بينما تعطى العينة الثانية ٣٩,٨% وتعطى العينة الثالثة ٤٠,٣ % وهكـــذا. وهـــنا نشـــير إلـــى الأوصـــاف الرقمية المأخوذة من العينة بأنها إحصـــاءات Statistics، وتكون فى الوقت ذاته تقديرية لمَعْالم المجتمع الأصلى Population parameters.

مــن خــلال مناقشــتنا لــتعريف الإحصـــاء وأساليبها المستخدمة فى العلوم الاجتماعية ذكرنا كلمات، مثــل البيانات Data، والمتغير The Variable، والمجتمع الأصلى، والعينة. فماذا نعنى بهذه الكلمات إحصائيًا؟

ثالثا: تعريف البيانات ومصادرها :

يقصـــد بكلمة بيانات Data فى الإحصاء الشكل الرقمى الذى يمثل خاصية أو ظاهـــرة مـــا. فمثلاً لو كان اهتمام أحد الباحثين دراسة درجة الوعى السياسى لدى عيـــنة من الشباب الجامعى، فإن البيانات التى تمثل درجة الوعى تقع فى شكل قيم علـــى مقـــياس للمشـــاركة السياسية، يعكس مدى اهتمام المبحوثين بمتابعة وتحليل القضـــايا السياســـية ومناقشتها مع الآخرين، ومدى المشاركة الإيجابية فى الأنشطة السياســـية مـــثل حضور الندوات السياسية العامة، والمشاركة بالتصويت فى العمل الانتخابى، وقراءة الخطب الانتخابية لمرشح معين ومناقشتها مع الآخرين، وارتداء شـــعار أو وضع لاصقة على السيارة لمرشح معين. ويصبح تعريف الإحصاء هنا، مجموعـــة الأســـاليب المســـتخدمة فى وصف درجة الوعى السياسى من معارف واتجاهـــات وممارسات للمبحوثين. ومن الأساليب التى يمكن أن يستخدمها الباحث فـــى هـــذا الوصـــف التوزيعات التكرارية، والرسومات البيانية، ومقاييس النزعة المركـــزية، ومقايـــيس التشـــتت. أما البيانات التى لا تعالج احصائيًا فيطلق عليها البيانات الخام Raw data.

مصادر البيانات :

تـــتعدد مصـــادر البـــيانات فقـــد تكون إحصاءات رسمية (جاهزة) من واقع السجلات والملفات، وقد يكون المصدر من الميدان من خلال أساليب متنوعة (مثل: صحيفة الاستبانه أو استمارة الحصر) يستخدمها الباحث فى جمع البيانات وفيما يلى نبذة عن كل مصدر.

١ – الإحصاءات الرسمية:

يشتمل هذا المصدر على جميع البيانات الإحصائية المدونة فى سجلات رسمية عـــن فـــترات زمنية ماضية ومحفوظة فى المؤسسات والهيئات كالجهاز المركزى للتعبـــئة العامـــة والإحصاء، السجل المدنى... إلخ، والمنظمات الدولية مثل منظمة العمل الدولية ILO، والبرنامج الإنمائى للأمم المتحدة (UNDP)... إلخ.

فـــإذا أردنــــا، على سبيل المثال، معرفة عدد سكان ريف وحضر مصر خلال الأعــوام ١٩٧٦، ١٩٨٦، ١٩٩٦، ٢٠٠٦ فيمكــن معرفة ذلك من خلال البيانات الإحصـــائية المدونـــة فى التعدادات العامة للسكان والصادرة عن الجهاز المركزى للتعبئة العامة والإحصاء خلال السنوات المذكورة.

٢- المصدر الميدانى:

يعـــتمد هـــذا المصـــدر على البيانات التى يقوم الباحث بجمعها من الميدان، مســـتخدمًا فـــى ذلـــك أسلوب أو أكثر من أساليب جمع البيانات الكمية، وتعتبر صـــحيفة الاســـتبانة أو الاستقصـــاء questionnaire أكـــثر الأساليب شيوعًا فى الاســـتخدام فـــى المســـوح الاجتماعية والإعلامية وتضم هذه الصحيفة قائمة من الأســـئلة (المفـــتوحة أو مغلقة النهاية) تغطى جوانب الظاهرة موضوع الدراسة وتطبق هذه الأداة من خلال المقابلة الشخصية أو هاتفيًا أو بإرسالها بالبريد أو من خـــلال البريد الإلكترونى e-mail. ولسهولة التحليل الإحصائى، يفضل أن تشتمل الاســـتبانة علـــى أســـئلة مغلقـــة النهاية أى تأتى الاستجابات لكل سؤال محددة. وكنموذج للأسئلة مغلقة النهاية، نورد فيما يلى عددًا منها على المبحوث أن يضع علامة (✓) أمام الإجابة الملائمة.

١- النوع :
(١) ذكر ( ) ☐
(٢) أنثى ( )

٢- الجنسية
(١) مصرى ( )
(٢) عربى ( ) ☐
(٣) أجنبى ( )

٣- الحالة التعليمية :
(١) دون المتوسط ( )
(٢) مؤهل متوسط ( ) ☐
(٣) جامعى أو ما يعادله ( )
(٤) ماجستير/ دكتوراه ( )

٤- السن:

(١) أقل من ٢٠ سنة ( )

(٢) ٢٠ - ٣٠ ( )

(٣) ٣٠-٤٠ ( )

(٤) ٤٠-٥٠ ( )

(٥) ٥٠-٦٠ ( )

(٦) ٦٠ سنة فأكثر ( )

٥- الحالة الاجتماعية :

(١) لم يسبق له الزواج ( )

(٢) متزوج ( )

(٣) مطلق ( )

(٤) أرمل ( )

رابعًا : أنواع المتغيرات :

المتغير المتصل Continuous Variable :

إن المتغـــير المتصل يأخذ أى قيمة متدرجة على المقياس المستخدم. مثال ذلك قـــياس درجـــات الحـــرارة باستخدام الترمومتر. فالمتغير يأخذ قيمة ما بين رقمين صـــحيحين بمعـــنى أن المتغير يمكن أن يأخذ أى قيمة بين ٣٦ درجة، ٣٧ درجة (٣٦,١، ٣٦,٢، ٣٦,٣، ٣٦,٤... الـــخ) وكذلـــك الحال فى مقاييس الطول والوزن ومستوى الذكاء وهذه المتغيرات يمكن أن يكون لها قيم كسرية. وقد يكون المطلوب التوصـــل إلى أقرب متوسط رقمى لعدد من الأرقام تختفى بها مشكلة بحثية ما مثل معـــرفة المتوســـط الحسابى لأوزان الأطفال حديثى الولادة. او متوسط الاستهلاك الشهرى لعينة من الأسر الحضرية.

المتغير المتقطع (البيانات المنفصلة) Discrete Variable:

عـــندما يـــأخذ المتغير قيمًا محددة ومنفصلة يطلق عليه المتغير المتقطع حيث يحتوى مداه على عدد محدود من القيم لا يمكن قياسها كميًا لأن كل جانب منها قائم بذاتـــه ومنفصـــلاً عن الآخر أى ليس له صلة بالجوانب الأخرى. فعدد الأولاد أو الأفـــراد فى الأسرة مثلاً لابد أن يكون رقمًا صحيحًا مثل ١، ٢، ٣، ٤... وهكذا. ومـــن أمـــثال المتغـــيرات المتقطعة، الجنس Sex (ذكور، إناث)، الحالة الزواجية Marital Status، ويشـــتمل علـــى أربع خواص هى (لم يسبق له الزواج، متزوج،

أرمـــل، مطلـــق). عدد أيام العمل فى أحد المصانع، عدد حوادث السيارات وعدد الكتب فى المكتبة.

المتغيرات المستقلة والتابعة :

أيضًا تنقسم المتغيرات إلى متغيرات تابعة Dependent Variables ومتغيرات مستقلة Independent Variable.

وتعـــرف المتغـــيرات المستقلة بأنها المتغيرات التى يستطيع الباحث أن يتحكم فيها والتى قد تؤثر طرديًا أو عكسيًا فى الظاهرة موضوع الدراسة (المتغير التابع). علـــى ســـبيل المثال، لو كان الهدف من بحث اجتماعى ما هو معرفة تأثير إدمان الســـيدات للمخدرات على الصحة الإنجابية لهن، ففى هذا البحث يتحكم الباحث فى الجرعة من حيث الكمية والنوع ثم يلاحظ التغيرات الصحية على أفراد العينة. فى هـــذه الحالـــة تمثل الجرعة المخدرة متغيرًا مستقلاً والصحة الإنجابية متغيرًا تابعًا. وبالنســـبة للبحوث الاجتماعية، تتعدد المتغيرات المستقلة التى تفسر حدوث ظاهرة مـــا. مثال ذلك العوامل التى تؤدى إلى انحراف الأحداث قد تكون التفكك الأسرى، أو مســـتوى تعلـــيم الوالديـــن، أو دخل الأسرة، أو خصائص شخصية للحدث، أو جماعة الرفاق.

ويعـــرف المتغـــير التابع بأنه تابع للمتغير المستقل أى كلما تتغير قيم المتغير المســـتقل تتغير تبعًا لذلك القيم المناظرة للمتغير التابع. ويستطيع الباحث أن يتعرف من خلال التغير فى قيم المتغير التابع كيفية ارتباطه بالمتغير المستقل (أهو ارتباط طردى أم ارتباط عكسى).

خامسًا: المجتمع الأصلى والعينة :

يشـــير المجـــتمع الأصلى Population إلى مجتمع البحث الذى يشتمل على جمـــيع الأفـــراد ويمكن أخذ عينات بحثية منه. وذلك لصعوبة إجراء البحوث على جمـــيع أفـــراد أو وحـــدات المجـــتمع لا سيما كبير الحجم. لذا فإن معظم البحوث الاجتماعـــية تعتمد على المسوح بالعينة. وتشير العينة Sample إلى شريحة ممثلة للمجـــتمع الأصـــلى أى تشـــتمل على جميع خصائصه. وتوجد طرق عديدة لأخذ العيـــنات مـــن مجـــتمع البحث، فعلى سبيل المثال يتم اختيار العينة العشوائية إما بالطـــريقة البســـيطة، أو باســـتخدام قوائم الأرقام العشوائية، أو بطريقة منتظمة أو طبقـــية سواء نسبية أو غير نسبية أو العينة العنقودية Cluster sample ونعنى بالعينة العشـــوائية Random Sample إعطـــاء فـــرص متساوية لجميع أفراد أو عناصر

المجتمع لكى يكونوا ضمن مفردات Subjects العينة المختارة. بمعنى أن يكون لكل عنصر داخل المجتمع الأصلى احتمال معلوم بوجوده فى العينة الممثلة لهذا المجتمع.

وقد يصعب على الباحث فى أحوال معينة اختيار عينة عشوائية نظرًا لعدم وجود إطار يشتمل على جميع عناصر المجتمع الأصلى أو أن يكون هذا المجتمع غير معلوم أو خفى Unknown or hidden society، مثل هذه الحالات يضطر الباحث إلى استخدام إحدى الطرق غير العشوائية فى العينة القصدية، العينة الحصصية Quota sample أو العينة المتاحة أو عينة كرة الثلج Snowball sample. ومن أهم عيوب الطرق غير العشوائية فى اختيار العينات، أن الباحث لا يستطيع تعميم generalization نتائج بحثه أى أنها تفتقر إلى الصدق الخارجى the external validity.

المعْلم Parameters والإحصاءات Statistics

يشير المعْلم إلى كل قيمة من القيم التى تتعلق بخصائص المجتمع الأصلى أما الخصائص المتعلقة بالعينة فيسمى كل منها تقديرًا estimate لقيمة تلك الخاصية فى المجتمع والتى على الأغلب تكون غير معلومة، ومن ثم يتم حساب تقديرها. وتشير الأوصاف الرقمية المسحوبة من العينة إلى أنها إحصاءات ولتكون فى الوقت ذاته تقديرية لمَعْالم المجتمع الأصلى Population وعادة تستخدم الحروف اللاتينية فى الإشارة إلى مَعْالم المجتمع الأصلى مثل (μ) يشير إلى المتوسط الحسابى، (σ) للانحراف المعيارى، ويستخدم الحروف الإنجليزية للتقديرات $\overline{X}$ متوسط حسابى، (S) للانحراف الميعارى للعينة (Blalock, 1972: 109-110; عبد الجبار توفيق، ١٩٨٣: ٢٣-٢٢٤).

مراحل البحث الإحصائى:

هناك عدة خطوات ينبغى على الدارس فى البحث الإحصائى أن يلتزم بها، وهى:

١- صياغة وتوجيه الأسئلة
٢- تحديد مجتمع البحث
٣- تصميم أداة/ أدوات جمع البيانات.
٤- تجميع البيانات من الميدان
٥- التحليل الاحصائى للبيانات
٦- تفسير نتائج البحث

سادسًا : مستويات القياس :

يقصد بالقياس أنه عملية تعبير عن الخصائص والمشاهدات بشكل كمى ووفقًا لقاعدة محددة.

ولعل أبسط أمثلة القياس نجدها فى الاختبارات المدرسية التى يتقدم لها الطلاب فى مختلف مراحل حياتهم الدراسية، حيث ترتبط الدرجة التى يحصل عليها كل طالب فى اختبار ما بمدى معرفته بالمادة التى يدرسها خلال فترة دراسية معينة، وكلما كانت درجة الطالب التى حصل عليها عالية دل ذلك على معرفة أكثر أو تحصيل أكبر لدى الطالب من هذه المادة كأن تكون هذه المادة هى مادة الكيمياء. ومن هذا المثال البسيط نجد أن خاصية التحصيل تعبر عنها الدرجة Score التى حصل عليها الطالب من الاختبار.

وتعتبر المقاييس التى تقيس المتغير التابع Dependent Variable واحدة من أكثر المقاييس أهمية عند تحديد الأساليب الإحصائية الملائمة التى تستخدم فى تحليل بيانات دراسة ميدانية معينة.

وتوجد أيضًا بعض المقاييس التى يمكن استخدامها فى قياس ظاهرة معينة بدقة عالية أو متناهية مثال ذلك المقاييس التى تستخدم فى قياس الأطوال والأوزان. كما توجد بعض المقاييس التى تفتقر إلى الدقة العالية وإن كانت تحقق قدرًا من الدلالة منها على سبيل المثال مقاييس مستويات القلق النفسى عند الأفراد أو درجة الاغتراب الاجتماعى لعينة من عمال الصناعة.

ومن ثم تنقسم أنواع المقاييس وفق مستوياتها من الأدنى للأعلى وما تتسم به من خصائص إلى:

مقاييس اسمية Nominal scales، ومقاييس رتبية Ordinal scales، مقاييس فاصلة Interval scales ومقاييس النسبة Ratio scales. وفيما يلى نبذة عن كل نوع من هذه المقاييس الأربعة.

المقاييس الاسمية :

يقوم هذا النوع من المقاييس بتصنيف الأفراد أو الأشياء أو المعلومات المتماثلة فى خاصية معينة فى مجموع أو فئة واحدة category. مثال ذلك إذا قمنا بتصنيف عدد من الأفراد وفق متغير الديانة : مسلم، مسيحى ويهودى. وقد نقوم أيضًا بعمل تصنيف آخر للفئات الثلاث على أساس الانتماء الحزبى (وطنى، الوفد، الغد، الناصرى والأحرار).

ومـــن خصـــائص هذا المقياس أنه لا يهتم بالتمييز أو التفضيل بين الفئات المخـــتلفة ففـــي المـــثال السابق لا نهتم بالتمييز بين الفئات الدينية على أساس الأهمـــية مـــثلاً، لا نقـــول إن المسلم أهم من المسيحى أو إن المسيحى أهم من الـــيهودى. كمـــا لا يوجـــد تداخـــل على أساس الديانة فالمجموعة كاملة تضم أفـــرادًا متماثلين فى نوع الديانة ومن ثم لا تتكرر مفردة فى أكثر من مجموعة (Blalock, 1972: 15-16; Hinkle, Wiersma and Jurs, 1979: 6).

المقاييس الرتبية Ordinal Scales:

فى المثال السابق، فضلاً عن تصنيف الأفراد إلى ثلاثة مذاهب دينية، يمكن أن ترتب تلك المجموعات الثلاث وفقًا لأهميتها أو لما تمتلكه كل منها من خاصية أو ســـمات معينة مشتركة. وقد نجد مثالاً أقرب للفهم فى الرياضيات عندما نميز بيـــن المقدارين (أ)، (ب) فنقول أن (أ) أكبر من (ب) ونأخذ الشكل الرياضى التالى:

أ > ب

وقـــد تكـــون أ > ب ولكـــن مقـــدار الفرق فى القيمة الدالة على التمييز بين أ، ب لـــيس مـــن خصائص المقياس الرتبى. ومن ثم فإن هذا المقياس ذو مستوى أعلـــى من المقياس الأسمى فى قياس الظاهرة أو الخواص. وتعتبر خاصية التمييز باســـتخدام علامـــات (>) أو (<) الخاصية الثانية إذا أخذنا فى الاعتبار أنه يشتمل على خاصية التصنيف وفق الترتيب.

وفـــى العلوم الاجتماعية نجد مثالاً لخاصية الترتيب دون الالتزام بالفروق عندما نصـــنف الأسر وفقًا للمكانة الاجتماعية – الاقتصادية Socioeconomic Status إلى طبقة عليا، ووسطى ودنيا.

وتشير الخاصية الثالثة إلى عدم تكرار نفس المفردة فى أكثر من مجموعة كما هو الحال فى المقياس الاسمى.

وأمـــا الخاصـــية الرابعة فهى الانتقالية. فلو فرضنا أن أ > ب وأ، ب > جـــ فيمكن القول أن أ > جـــ. ولكن من المنظور الترتيبى.

ويجـــدر التـــنويه إلـــى ضرورة ملاحظة أن المستوى الرتبى للمقياس لا يهتم بالفروق بين العناصر أو الخواص. ومن ثم لا نستطيع أن نستخدم مع هذا المقياس العملـــيات الحسابية كالطرح والقسمة والضرب والجمع كما أننا لا يمكن استخدامها أيضًا مع المقياس الاسمى.

المقاييس الفاصلة:

مـــن خصـــائص المقياس الفاصل Interval Scale بالإضافة للخصائص التى ذكرناها فى المقياسين السابقين، توحيد نوع وحدة القياس، فلا يمكن أن نقيس الفرق بين درجتين من الحرارة إحداهما بالفهرنهايت والأخرى بالدرجة المئوية، بل يكون الفرق بين درجتين حراريتين مثل ٣٨ درجة مئوية، ٣٠ درجة مئوية أى من نفس جنس وحدة القياس.

مـــن جهـــة أخـــرى، إذا قلنا، إنه توجد وحدات قياسية للمقياس الفاصل، ففى العلـــوم الاجتماعية قد يتعذر تحقيق ذلك، فمثلاً لا توجد وحدات قياسية أو معيارية لقـــياس الســـلطة، أو الهيبة الاجتماعية التى نجدها متكررة دائمًا فى الموضوعات الاجتماعية.

والخاصية الثانية للمقاييس الفاصلة إمكانية استخدام العمليات الحسابية المختلفة من جمع وطرح وضرب وقسمة للدرجات فى عمليات تحليل البيانات. فمثلاً يمكن إضافة دخل الزوجة إلى دخل الزوج أو إلى دخل باقى أفراد الأسرة.

أما الخاصية الثالثة للمقياس الفئوى فهى أنه يهتم بخاصية تساوى الفروق بين المســـتويات المخـــتلفة مـــثال ذلـــك تقســـيم الدرجة الواحدة على مقياس الحرارة (الترمومـــتر) إلى خمسة أقسام يمثل كل جزء منها ٠,٢ من الدرجة. ويطلق على هذا النـــوع من المقاييس مقياس الفئات المتساوية Equal intervals كما لا يشتمل هذا المستوى من القياس على نقطة الصفر المطلق وإنما الصفر يعتبر نسبيًا.

القياس النسبى:

يعتـــبر القـــياس النسبى Ratio من أرقى مستويات القياس ويشتمل على جميع الخصـــائص السابقة. فضلاً عن وجود الصفر المطلق الذى يعنى غياب الخاصية. والقياس النسبى ليس محور اهتمامنا فى البحوث الاجتماعية.

سابعًا: التوزيع التكرارى :

يُعَرَّف التوزيع التكرارى Frequency Distribution بأنه عملية ترتيب الأرقام فـــى صـــورة تعطـــى عدد مرات حدوث الرقم أو الصفة أو ما نسميها بالتكرارات. بالإضـــافة إلـــى أن التوزيع التكرارى ينظم البيانات تنظيميًا نمطيًا كما يعطى الباحث بعـــض الـــدلالات عـــن طبيعة البيانات التى بين يديه. من جهة أخرى فإن التوزيع الـــتكرارى لا يعطـــى الباحث أى معلومة حول تلك البيانات بل يمثل مقدمة لإجراء التحليل الاحصائى واختبار الفروض.

خصائص التوزيع التكرارى :

توجد أربع خصائص تحدد المفهوم السابق للتوزيع التكرارى وهى:

الوضع المركزى Central Location:

ونعنى به قيمة محددة تتوسط باقى قيم التوزيع. وتستخدم المتوسطات كمقاييس إحصـــائية فـــى تقديـــر تلك القيمة والمقاييس للمتوسطات هى المتوسط الحسابى، المـــنوال، والوســـيط، وتعـــرف بمقاييس النزعة المركزية والتى سنعرض لها فى الفصـــل الـــرابع. ونظرًا لاختلاف طريقة الحساب من مقياس إلى آخر فى حساب المتوسطات، فإن القيمة الوسطى سوف تختلف قيمتها تبعًا لذلك.

التشتت أو التباين:

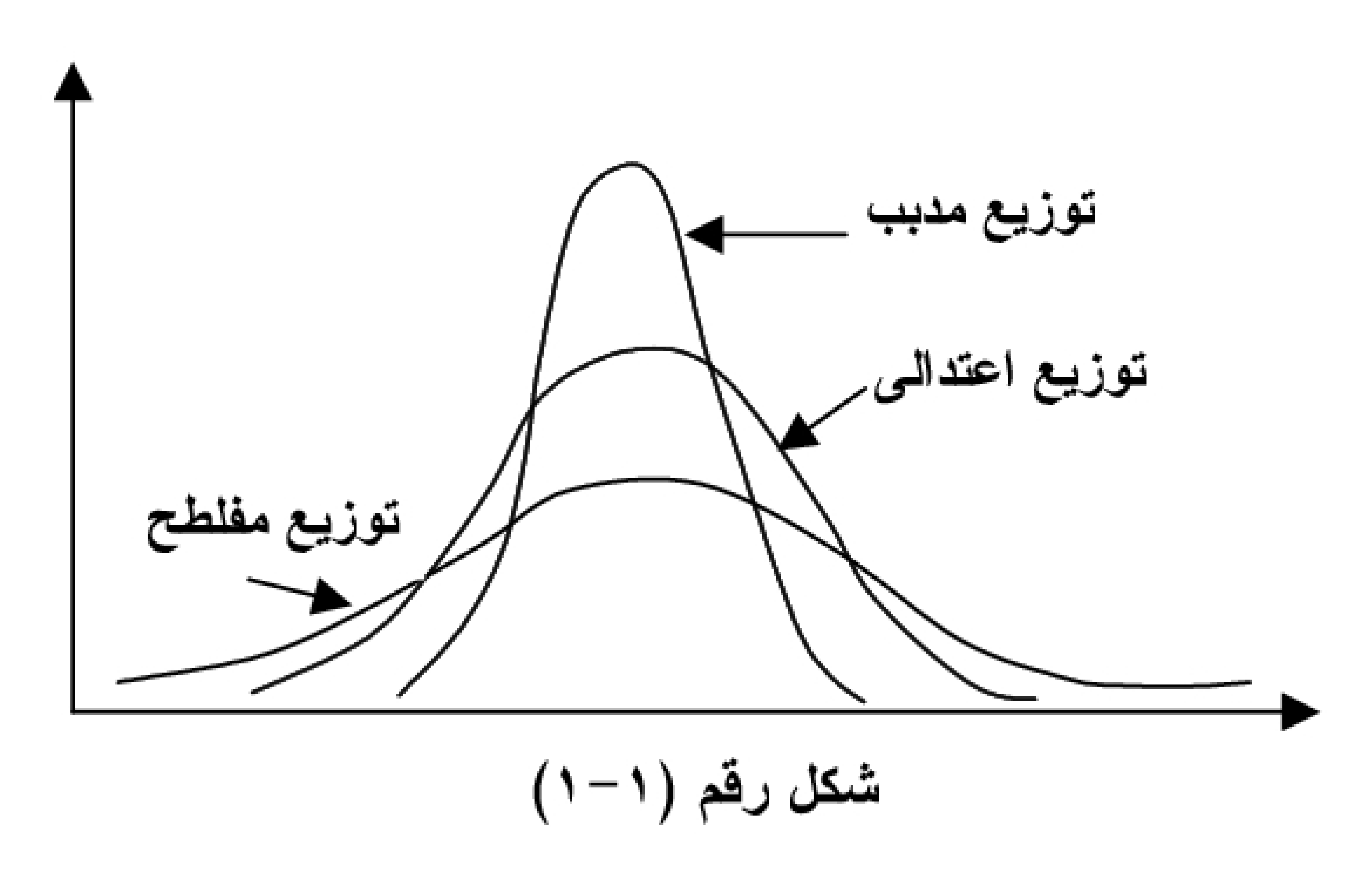
ويقصـــد بـــه درجة انتشار أو تشتت القراءات المختلفة للظاهرة حول قيمتها الوســـطى. وأنـــه كلمـــا ازداد تجمع المفردات حول تلك القيمة الوسطى قل تشتتها وبالتالى يقال للتكرارات المشاهدة عن تلك الظاهرة إنها أكثر تجانسًا.

الالتواء:

وهـــى خاصـــية تعـــنى ابـــتعاد التوزيع التكرارى عن التماثل ويمكن قياس الالتواء Skewness بعدة طرق من بينها المنوال والوسيط، والربيعين. وقد يطلق على التوزيع التكرارى أنه موجب الالتواء Positively Skewed إذا كانت معظم التكرارات متجمعة عند القيم الكبرى ناحية اليمين. أما إذا كانت معظم التكرارات مـــتجمعة عند القيم الصغرى من ناحية اليسار من المنحنى التكرارى فيطلق على التوزيع أنه سالب الالتواء Negatively Skewed.

التفلطح:

تعتـــبر أهمـــية هـــذه الخاصية فى أنها تحدد مدى اختلاف التوزيع التكرارى للظاهـــرة عـــن التماثل للتوزيع الاعتدالى. بالإضافة إلى خاصية الالتواء. فالتفلطح Kurtosis ينـــبأ عما إذا كان للتوزيع التكرارى للقيم قمة عليا حادة أو قمة عريضة مســـطحة. ففى الشكل رقم (١-١) نجد لدينا ثلاثة منحنيات تكرارية أوسطها توزيع اعـــتدالى Normal Distribution أما المنحنيان الآخران فالأول له قمة مدببة تعلو قمة المنحنى المدبب، أما المنحنى الثالث والذى تقل قمته عن قمة التوزيع الاعتدالى وتـــأخذ القمـــة شـــكلاً أكثر استوائية وتفلطحًا تبعًا لتشتت القيم التى يشملها التوزيع ويطلـــق عليه منحنى تكرارى مفلطح. ومن ثم يمكن أن نميز بين التوزيعين نسبيًا بالتوزيع الاعتدالى كما يلى:



شكل رقم (١-١)

أ - **التوزيع المدبب:** هو التوزيع الذى تكون تكراراته المركزية أكبر من التكرارات فى التوزيع الاعتدالى كما تزداد فيه خاصية تجمع التكرارات بالقرب من الفئات الوسطى.

ب- **التوزيع المفلطح:** هو التوزيع الذى تقل تكراراته المركزية عن التوزيع الاعتدالى كما تنتشر تكراراته على مدى أكبر حول الفئات الوسطى ويمكن قياس التدبب والتفلطح على أساس الفرق بين مدى ارتفاع قمة المنحنى الاعتدالى عن المحور الأفقى وكلٍ من قمتى التوزيع المدبب والمفلطح على التوالى.

حساب معامل التفلطح :

يمكن حساب قيمة معامل التفلطح للتوزيعات التكرارية من المعادلة التالية:

$$\text{معامل التفلطح} = \frac{\text{نصف الفرق بين الربيعين الأعلى والأدنى}}{\text{الفرق بين المئينين التسعون والعاشر}}$$

$$= \frac{\text{الانحراف الربيعى}}{\text{المئين التسعون} - \text{المئين العاشر}}$$

مثال:

تم حساب القيم للمقاييس الأربعة فى المعادلة السابقة لقياس معامل التفلطح لإحدى التوزيعات التكرارية. فكانت تلك القيم كالآتى:

الربيع الأعلـــى = ٧٢,٤٥

الربيع الأدنـــى = ٥٦,٥

المئين التسعون = ٧٩,٤٥

المئين العاشر = ٤٤,٧٥

والمطلوب حساب معامل التفلطح لتلك التوزيعات.

الحـــــل

$$\text{معامل التفلطح} = \frac{\frac{\text{٧٢,٤٥} - \text{٥٦,٥٥}}{\text{٢}}}{\text{٧٩,٤٥} - \text{٤٤,٧٥}} = \frac{\text{٧,٩٥}}{\text{٣٤,٧٠}} = \text{٠,٢٢٩}$$

وباســتخدام الخصــائص الأربــع الســابقة للتوزيع التكرارى يمكن أن نقسم التوزيعات التكرارية بشكل عام إلى نوعين أساسيين هما :

أ – توزيعات تكرارية اعتدالية.

ب– توزيعات تكرارية غير متماثلة Asymmetrical distribution:

وتعــرف بأنها التوزيعات التى تتناقص أو تزداد فيها التكرارات للقيم بصورة غــير اعتدالــية أو غــير منتظمة على جانبى المحور الرأسى المقام عند منتصف التوزيع والذى يقطع المحور السينى أو المحور الأفقى فى التمثيل البيانى.

المفاهيم الأساسية Key Concepts

الإحصاء:

مجموعـــة مـــن الأســـاليب العلمـــية المســـتخدمة فى جمع البيانات وتنظيمها وتلخيصها وتحليلها بغرض الوصول إلى قوانين وقرارات.

الإحصاء الوصفى:

فـــرع مـــن الإحصـــاء يختص بتلخيص ووصف توزيع متغير واحد وبقياس العلاقة بين متغيرين أو أكثر.

الإحصاء الاستدلالى:

فـــرع من الإحصاء يختص بعمل تعميمات للمجتمع الأصلى من خلال ما يتم أخذه من عينات ممثلة له.

العينـــة:

شـــريحة مختارة بعناية من المجتمع الأصلى ويتم الاختيار بعدة طرق، وعلى الباحـــث أن يســـتخدم الطـــريقة الأكـــثر ملاءمة لأهداف بحثه وفى ضوء البيانات الإحصائية المتاحة عن خصائص المجتمع الأصلى.

المتغير:

أيـــة خاصـــية يمكـــن أن تتغير قيمها من مفردة لأخرى داخل العينة البحثية. (كالطول، الوزن، العمر، الحالة الزواجية).

الثابت:

أيـــة خاصـــية يفترض احتفاظها بقيمة ثابتة لجميع الأفراد داخل مجموعة ما موضوع الدراسة.

المتغير المتصل Continuous Variable:

هو المتغير الذى يأخذ أى قيمة متدرجة على المقياس المستخدم.

المتغير المتقطع Discerte Variable:

هـــو المتغير الذى يأخذ قيمًا محددة أو هو الذى يحتوى مداه على عدد محدود أو لانهائى من القيم بشرط أن يكون لكل منها قيمة محددة يمكن ترتيبها.

المقياس الاسمى Nominal Scale:

هو عملية تصنيف الموضوعات المختلفة إلى فئات تعتمد على سمات المتغير المحددة أو خواصه.

المقياس الرتبى Ordinal Scale:

ويتميز عن المقياس الاسمى بأنه يحتوى على ترتيب منطقى للفئات فضلاً عن اكتسابه صفات هذا المقياس.

المقاييس الفاصلة Interval Scale:

وهـــى مقاييس تتضمن خصائص المقياسين السابقين، بالإضافة إلى أن الفروق بيـــن الفـــئات المختلفة متساوية، مع تواجد وحدة القياس، ويمكن استخدام العمليات الحسابية (من ضرب وقسمة وجمع وطرح) فى تحليل البيانات.

التوزيع التكرارى Frequency Distribution:

هـــو عملـــية ترتيـــب الأرقـــام فى صور تعطى عدد مرات تكرار الرقم فى المجموعة.

تماريـــن

١- أكمل العبارات الآتية :

(أ) .... هـــى مجموعـــة الأساليب الفنية المستخدمة من جانب علماء العلوم الاجتماعية بقصد تنظيم ومناقشة البيانات بهدف اختبار النظريات والإجابة على أسئلة البحث.

(ب) تعتـــبر الإحصــــاء وتطبيق الأساليب الإحصائية من الأمور الحيوية من أجل .... البحث.

(ج) يعــرف .... بأنه خاصية قد تكتسب قيمًا مختلفة من حالة إلى حالة أخرى.

(د) علـــى الباحـــث الاجــتماعى الـــذى يريد وصف وتلخيص وتوزيع ظاهرة أو خاصـــية فـــى مجتمع أن يستخدم الأساليب الإحصائية لأحد فرعى الإحصاء المسمى ....

(هـ) يســـتخدم الباحث الإحصاء .... عندما يريد فهم العلاقة بين متغيرين أو أكثر.

(و) يمـــثل الإحصـــاء الاســـتدلالى مجموعة الأساليب الإحصائية التى يستخدمها الباحثون عندما يرغبونه فى التعميم من .... إلى ....

(ز) عـــندما تنشـــر صـــحيفة (المساء) المصرية استفتاءً يشير إلى أن ٦٥% من المصـــريين يشـــاهدون كـــرة القـــدم، فإنهـــا اســـتعانت بأســـاليب الإحصاء ....

٢- فـــيما يلى عدد من الاختيارات إحداها يمثل الإجابة الصحيحة على كل سؤال مـــن الأســـئلة الآتية. ضع علامة (✓) أمام الاختيار الذى تراه إجابة صحيحة على السؤال.

- **إن قطر القمر بالأميال يمثل:**

(أ) مجتمع أصلى.

(ب) ثابت.

(ج) إحصائى.

(د) مَعْلَم a parameter.

٣- أجـــرى مسح اجتماعى على عينة عشوائية قوامها ٥٠٠ شابًا فى حى صغير بمديـــنة القاهـــرة بهدف معرفة اتجاهاتهم نحو سياسة الحكومة الحالية مقارنة

بالحكومـــة الســـابقة فـــيما يختص بالاهتمام بقضايا الشباب. فأى اختيار من الاختيارات الآتية يعتبر المتغير التابع:

(أ) حجم الحى الصغير فى مدينة القاهرة.
(ب) عدد المبحوثين فى العينة.
(ج) الاتجاهات نحو سياسة الحكومة المصرية.
(د) الاختيارات الثلاثة السابقة ليس من بينها المتغير التابع.

٤- فـــى محاولـــة لتقدير سن الطالبات فى قسم الاجتماع بكلية البنات جامعة عين شـــمس، قـــام أســـتاذ مادة الإحصاء بأخذ متوسط العمر للطالبات داخل قاعة المحاضرات. فهل يمثل هذا المتوسط.

(أ) إحصاء.
(ب) عينة.
(ج) مجتمع أصلى.
(د) مَعْلَمْ.

٥- هل الإحصاء الاستدلالى :

(أ) هو الإحصاء الوصفى.
(ب) يستخدم بعض أساليب الإحصاء الوصفى.
(ج) يتيح للباحث التعميم للمجتمع الأصلى من خلال العينة.
(د) الاختياران الثانى والثالث السابقان معًا.

٦- أراد أســـتاذ مادة الإحصاء أن يقارن بين أسلوبين للتدريس لمادة الإحصاء فى فصـــلين دراســـيين مختلفين. واعتمد فى المقارنة على الدرجات التى تحصل عليها الطالبات فى كل فصل منهما فى الاختبار النهائى لمادة الإحصاء.

كـــم عـــدد المتغيرات التى سيقوم الأستاذ باختبارها؟ وما هو المتغير المستقل والمتغير التابع فى هذه الدراسة.

٧- أجـــرى مسح قومى على عينة عشوائية قوامها ١٦٠٠ وحدة معيشية فى حى شـــبرا بمدينة القاهرة بهدف التعرف على الموافقة أو عدم الموافقة من جانب الأفـــراد على سياسة الحكومة المصرية فى تخفيض أسعار السلع الإستهلاكية. وكشفت نتائج المسح عن أن ٥٥% من الأفراد يوافقون على سياسة الحكومة. فى هذا المسح. حدد ما يلى:

(أ) المجتمع الأصلى.
(ب) العينة.

(ج) نوع الأسلوب الإحصائى (استدلالى أم وصفى) الذى تم استخدامه فى هذا المسح.

٨- فـــى السؤال رقم (٤)، تمثل الطالبات داخل قاعة المحاضرات بالنسبة لأستاذ مادة الإحصاء.

(أ) إحصاء.

(ب) عينة.

(ج) مجتمع أصلى.

(د) مَعْلَمْ.

٩- أراد فريق من علماء الاجتماع السياسى أن يعرف الانتماءات الحزبية لطلاب جامعة القاهرة وذلك من خلال إجراء المقابلات معهم أثناء فترة تسجيلهم للعام الدراســى الجديد. وكشفت المقابلة عن أن ٤٥% من الطلاب ينتمون للحزب الوطـــنى، ٣٠% لحـــزب الوفـــد، ٢٥% لحزب العمل. ما هى العينة فى هذا المـــثال؟ ومـــا هـــو المجتمع الأصلى؟ وأى الأساليب الإحصائية (وصفية أم استدلالية) ينتمى إليها هذا المثال؟



الفصل الثانى

تبويب البيانات

مقدمة

١- تبويب البيانات.

٢- الجداول التكرارية للبيانات الكمية.

٣- الجداول التكرارية البسيطة للبيانات الكيفية.

٤- الجداول التكرارية المزدوجة للبيانات الكيفية.

الفصل الثانى
تبويب البيانات

مقدمة :

من الشائع فى مجال البحوث الاجتماعية توافر كم من البيانات الإحصائية التى يحصل عليها الباحث باستخدام أدوات جمع البيانات المناسبة، وعادة تتمثل تلك البيانات فى شكل أرقام، تعتبر قياسًا للمتغيرات مجال الدراسة، ولما كانت تلك الأرقام تفتقر إلى الترتيب والتصنيف يطلق عليها البيانات الأولية أو البيانات الخام Raw Data، وهى البيانات التى لم تعالج إحصائيًا.

وتعرف البيانات الإحصائية أنها كمية من المعلومات على هيئة أرقام وأن تلك الأرقام إما أن تكون أرقامًا صحيحة Integers مثل ١٠، ٢٠، ٣٠ وهكذا أو تكون أرقامًا عشرية أو حقيقية Real Numbers مثل ٨,٥، ١٠,٢٥، ١٥,٥ وهكذا .. ويتوقف حجم البيانات الخام على حجم المجتمع الأصلى، فكلما إزداد حجم هذا المجتمع يتوقع مزيدًا من الأرقام والتى يصعب مع كثرتها وعدم تصنيفها تفهم أو قياس متغير أو أكثر موضوع الدراسة، ومن ثم كان من الضرورى أن يقوم الباحث بتصنيف وتبويب تلك البيانات بالشكل أو بالأسلوب الذى يخدم هدف البحث بشكل جيد، من دراسة المتغيرات أو استنباط نوعية العلاقات أو المعلومات المهمة التى تتعلق بتلك المتغيرات. فمثلاً لو أجرينا اختبارًا لقياس القدرات لعدد (١٢٠) طالبًا، فإنه يلزم استخدام (١٢٠) رقمًا مناظرًا بواقع رقم لكل طالب. فلو افترضنا أن عدد الطلبة يصل إلى خمسة أو ستة أضعاف هذا العدد. فمن المؤكد أننا سوف نواجه صعوبات كثيرة فى قياس القدرة ما لم نستخدم أداة احصائية أو أكثر لتنظيم البيانات الخام بهذا الحجم الكبير. ولعل أبسط الطرق الإحصائية لتنظيم وتلخيص البيانات هى طرق التوزيع التكرارى Frequency Distribution.

ويعرف التوزيع التكرارى بأنه عملية ترتيب الأرقام فى صورة تعطى عدد مرات تكرار الرقم أو الصفة. فالتوزيع التكرارى هو طريقة لتصنيف البيانات وترتيبها وتقسيمها تقسيمًا يساعد الباحث على إدراك ما بينها من علاقات كما أنه لا يعطى للباحث أى معلومة حول تلك البيانات بل يمثل مقدمة لإجراء التحليل الإحصائى واختبار الفروض.

ويهــدف الفصل إلى أن يتدرب الطالب على عمل جدول لتفريغ البيانات الخام (الكمــية والكيفــية). وأن يعــرف كيفــية عمل جداول التكرار التجمعى والجداول التكرارية البسيطة والمزدوجة)، وكيفية حساب النسب المئوية.

تبويب البيانات Tabulation :

أوضــحنا فــيما سبق أهمية التوزيع التكرارى كوسيلة لتصنيف البيانات الخام حــول الظاهــرة المطلــوب دراستها وتجميع بيانات يتم تحويلها إلى بيانات رقمية تمهيدًا لتحليلها إحصائيًا. وقلنا إن البيانات الرقمية قد تكون كثيرة وذات قيم مختلفة سواء كانت متقاربة فى قيمتها أو متباينة. ومن ثم لا يجد الباحث مفرًا من ضرورة تصنيف تلك البيانات ومحاولة وضعها فى شكل جداول تمكن من عرضها بصورة تلخص معالمها وتساعد على استخلاص النتائج منها وهذا ما نعنيه إحصائيًا بعملية التبويب Tabulation.

ويــتم التبويــب للبيانات عادة على أساس كمى أو كيفى (نوعى) أو جغرافى أو زمنى. كما يمكن أن يتم التبويب بأسلوب المزج بين هذه الأسس. وتصنف المتغيرات إلى متغيرات كمية (مثل الدخل، العمر، الوزن، الطول، درجة الحرارة)، أو متغيرات كيفــية أو اســمية (مثل الحالة الزواجية، الجنس، المهنة والجنسية). فبالنسبة للمتغير الكمى، تحمل القيمة معنى كميًا ويتم ترتيب مفردات البحث حسب الخاصية موضوع الدراســة، كالعمر مثلاً، فى هذه الحالة، يتم الترتيب من الأكبر سنًا إلى الأصغر سنًا أما فى حالة المتغير الكيفى، فإن القيمة لا تشير إلى مقدار الخاصية، بل تعبر إما عن وجــود تلــك الخاصية أو غيابها. وقد يعطى الباحث قيمًا رقمية لصفات المتغير كأن يعطى رقم (١) ليدل على أن المبحوث أعزب ورقم (٢) للمتزوج، إلا أن هذه الأرقام لا تشــير إلــى مقــدار الخاصية أو أن رقم (٢) أكبر من رقم (١). بل تستخدم هذه الأرقام لغرض تصنيف صفات المتغير وليس لها أى دلالة رقمية.

الجداول التكرارية للبيانات الكمية (الرقمية) :

يمكــن للباحث أثناء قيامه بتصنيف البيانات الكمية أن يختار الفئات التى يحددها لنفسه ولكن بشرط أن تسهل عدد الفئات ومداها من إدراك ما بين البيانات الإحصائية مــن علاقــات ومــا لهــا من صفات ودلالات. ونعنى بالفئة تلك المجموعة الرقمية الجزئــية داخــل مدى البيانات التكرارية بحيث تشتمل على عدد من القيم المتقاربة. وبعد ذلك يقوم الباحث بحصر العدد الذى يقع داخل حدى كل فئة (الحد الأدنى والحد

الأعلـــى) ويســـمى تكـــرارًا ويرمز له بالرمـــز (ك). وبعد تكرار ذلك العمل لجميع البـــيانات داخل المدى المحدد لها بأعلى قيمة وأدنى قيمة من خلال ترتيب تنازلى أو تصـــاعدى فإن الباحث سيحصل فى النهاية على ما يسمى بالجدول التكرارى والذى يتضـــمن عـــددًا من الفئات وتكراراتها والذى يصبح أساس دراسة الظاهرة موضوع البحـــث. ويطلق على أبسط الطرق لتنظيم البيانات الإحصائية بالمصفوفة الرقمية أو العدديـــة Array وهـــذه الطـــريقة تصـــلح إذا كانت القيم صغيرة. أما إذا كانت القيم بالآلاف ومضاعفاتها فيصعب تمامًا استخدام المصفوفة ومن هنا تبرز أهمية استخدام الجدول التكرارى وأيضًا المنحنيات التكرارية Frequency Curves.

الجداول التكرارية البسيطة :

تعد الجداول التكرارية Frequency Tables إحدى وسائل عرض البيانات فى شـــكل يسهل معه التحليل والوصول للنتائج التى تتطلبها الدراسات، حيث إن تجميع البـــيانات فى شكل فئات ذات تكرارات تقلل من حجم توزيع البيانات ومن ثم يسهل استيعاب القيم الواردة بالجدول.

ولتوضيح طريقة عمل الجدول التكرارى نأخذ المثال التالى :

مثال :

فـــيما يلـــى بـــيان بـــالدخل الأسبوعى لعينة من العمال قوامها (٤٠) عاملاً.
المطلوب: تكوين جدول التوزيع التكرارى للدخل الأسبوعى لهم (بالجنيه المصرى).

٣٥ – ١٤ – ٣٢ – ٢٥ – ٢٨ – ١٨ – ٤٢ – ٢٣
١٨ – ٢٢ – ٢٦ – ١٩ – ٣١ – ٤٤ – ٣٢ – ١١
٢٦ – ٣٧ – ٢٢ – ٣٢ – ٢٦ – ٢٣ – ٣٨ – ٢٦
٢٣ – ٤٢ – ٢١ – ٢٦ – ١٧ – ١٤ – ٣٤ – ٣٥
٢٧ – ٤١ – ٣٩ – ٤١ – ١٤ – ٣٧ – ٢٥ – ٤٨.

خطوات الحل :

١– ترتيب القيم تصاعديًا :

١١ – ١٤ – ١٤ – ١٤ – ١٧ – ١٨ – ١٨ – ١٩
٢١ – ٢٢ – ٢٢ – ٢٣ – ٢٣ – ٢٣ – ٢٥ – ٢٥
٢٦ – ٢٦ – ٢٦ – ٢٦ – ٢٦ – ٢٧ – ٢٨ – ٣١
٣٢ – ٣٢ – ٣٢ – ٣٤ – ٣٥ – ٣٥ – ٣٧ – ٣٧
٣٨ – ٣٩ – ٤١ – ٤١ – ٤٢ – ٤٢ – ٤٤ – ٤٨

٢- تحديد أعلى درجة وأقل درجة.

٣- حساب المدى على النحو التالى : المدى = أكبر قيمة - أقل قيمة + ١

... المدى = ٤٨ - ١١ = ٣٧ + ١ = ٣٨

٤- تحديد طول الفئة من العلاقة الآتية :

$$\text{طول الفئة} = \frac{\text{المدى}}{\text{عدد الفئات المقترح}}$$

$$\text{طول الفئة} = \frac{٣٨}{٨} = ٤,٧٥ = ٥ \text{ تقريباً}$$

(حيث عدد الفئات المقترح = ٨)

ولا توجد قاعدة محددة لعدد الفئات فلا يكون عددها صغيرًا جدًا بحيث تضيع معــالم الظاهــرة ولا يكــون عددها كبيرًا بحيث تكون هناك خانات صفرية. ويقــترح بعض العلماء أن تتراوح عدد الفئات ما بين ٥ إلى ١٥. وتزيد عن هذا العدد فى التعدادات السكانية.

٥- عمــل جــدول لتفريغ البيانات مكون من ثلاثة أعمدة يشتمل العمود الأول قيم المتغــير (الدخــل) ونرمز له بالرمز (س) والعمود الثانى يخصص للعلامات والعمــود الثالث يشتمل على التكرارات والتى نرمز لها بالرمز (ك). ونسجل فــى العمــود الأول القيم من الأدنى للأعلى بنفس الترتيب دون ترك أى قيمة حتى إذا كانت غير موجودة فى البيانات. ونضع فى عمود العلامات خط مائل (/) أمام كل قيمة تتكرر بحيث يصبح عدد العلامات مساويًا لعدد مرات تكرار القيمة مع شطب القيمة من الكشف الأصلى وعندما تصبح عدد العلامات أربع (////) نضــع العلامة الخامسة بشكل مائل بحيث تكون حزمة (////)، كما هو موضــح فــى الجدول رقم (٢-١). أما العمود الأخير فترصد فيه التكرارات مساوية لعدد العلامات.

جدول رقم (٢-١)
توزيع البيانات المتعلقة بالدخل (بيانات كمية)

الدخل	العلامات	ك
١١	/	١
١٤	///	٣
١٧	/	١
١٨	//	٢
١٩	/	١
٢١	/	١
٢٢	//	٢
٢٣	///	٣
٢٥	//	٢
٢٦	~~////~~	٥
٢٧	/	١
٢٨	/	١
٣١	/	١
٣٢	///	٣
٣٤	/	١
٣٥	//	٢
٣٧	//	٢
٣٨	/	١
٣٩	/	١
٤١	//	٢
٤٢	//	٢
٤٤	/	١
٤٨	/	١
		مجـ = ٤٠

جدول رقم (٢-٢)

جدول تفريغ البيانات المتعلقة بالدخل

ف	العلامات	ك
١٠-	////	٤
١٥-	////	٤
٢٠-	~~////~~ /	٦
٢٥-	~~////~~ ////	٩
٣٠-	~~////~~	٥
٣٥-	~~////~~ /	٦
٤٠-	~~////~~	٥
٤٥-٥٠	/	١
مجـــ		٤٠

جدول رقم (٢-٣)

جدول تكرارى لتوزيع العمال حسب الدخل

فئات الدخل	ك	%
١٠-	٤	١٠,٠
١٥-	٤	١٠,٠
٢٠-	٦	١٥,٠
٢٥-	٩	٢٢,٥
٣٠-	٥	١٢,٥
٣٥-	٦	١٥,٠
٤٠-	٥	١٢,٥
٤٥-٥٠	١	٢,٥
مجـــ	٤٠	١٠٠,٠

نلاحـــظ فـــى الجـــدول (٢-١) أننا رتبنا الأرقام بفاصل واحد فقط إلا أنه فى بعـــض الحالات يمكن إعادة تنظيم الأرقام على شكل فئات بفاصل أكبر من الواحد

كـــأن يكون الفاصل مقداره الحسابى خمسة وهذا الفاصل يسمى طول الفئة Class Interval وهـــذا من شأنه أن يعطى مقارنة للبيانات بصورة أكثر سهولة. ولكل فئة حدان (الحد الأدنى والحد الأعلى).

ففـــى المثال السابق يمكن إعادة ترتيب مجاميع الأرقام لتصبح كما هو موضح فـــى الجدول رقم (٢-٢) بفاصل خمسة بين كل فئة. ونجد أننا أخذنا فئات متساوية المـــدى أو الطول وهو (٥) أو بمعنى آخر يمكن القول إن الفرق بين مركز كل فئة ومركز الفئة التى يليها يكون مساويًا ومقداره (٥) درجات.

$$\text{ومركز الفئة} = \frac{\text{الحد الأدنى للفئة + الحد الأعلى للفئة}}{٢}$$

يـــراعى عـــند تصميم الفئات أن يشمل الحد الأدنى للفئة الأولى أقل قيمة فى التوزيع وأن يشمل الحد الأعلى للفئة الأخيرة أعلى قيمة فى التوزيع كما هو واضح فى المثال السابق.

طرق كتابة الفئات :

أ – فى حالة البيانات المتقطعة :

يقـــوم الباحـــث بتحديد الحدين الرقمين الأدنى والأعلى لكل فئة بشكل واضح فضلاً عن كتابتها على هذا النحو فى توزيعات عدد أفراد الأسرة:

١–٣

٤–٦

ب– فى حالة البيانات المتصلة :

يقـــوم الباحـــث بتوضيح الحد الأعلى لكل فئة بينما يتم تحديد الحد الأدنى بها ضـــمنيًا مع تجنب حدوث تداخل بين نهاية فئة وبداية الفئة التى تليها هكذا، حتى لا يتكرر ازدواج لعدد من المفردات.

مثال :

فى توزيع المفردات وفقًا للفئات العمرية تكتب على النحو التالى:

٢٠ –

٣٠ –

٤٠ –

كما يمكن كتابة الحدود الدنيا فقط للفئات على الشكل التالى :

١٠ –
٢٠ –
٣٠ –

جدول التوزيع التكرارى النسبى والمئوى:

يقصد بالتكرار النسبى للفئة، كما سبق وأوضحنا، هو تكرارها بالنسبة إلى التكرار الكلى لجميع الفئات أو بالأحرى هو حاصل قسمة قيمة تكرارات الفئة على مجموع قيم التكرارات لجميع الفئات الموجودة بالجدول، وعادة يعبر عن هذا الناتج فى شكل نسبة مئوية أى يتم ضرب الناتج فى ١٠٠ للحصول على تلك النسبة (%) ويطلق على التكرار فى هذه الحالة بالتكرار المئوى.

والغرض الأساسى من عمل جداول التكرار المئوى هو احتياج الباحث فى بعض الأحيان إلى معرفة نسبة كل تكرار إلى التكرار الكلى. فمثلاً يحتاج الباحث إلى هذا التكرار المئوى عندما يريد معرفة أكثر الأسر الريفية إنجابًا للأطفال، أوأعلى نسبة وفيات بين الأطفال بسبب مرض معين، أو أعلى الفئات الوظيفية تقاضيًا للرواتب الشهرية فى شركة ما.

مثال :

فيما يلى بيانات الأجور الأسبوعية لعدد (٢٥) عاملاً باليومية فى إحدى الشركات الصناعية، وأرادت إدارة الشركة معرفة أكثر هؤلاء العمال انتظامًا ومواظبة فى العمل وذلك من خلال :

١- أن أعلى الأجور تدل على الانتظام فى العمل حيث لا يتقاضى عامل اليومية أجرًا إذا غاب عن العمل فى يوم من الأيام.
٢- إن أقل أجر يدل على عدم المواظبة فى الحضور للعمل.
٣- ففى هذا المثال يمكن استخدام جداول التكرار النسبى والمئوى، حيث أن النسبة المئوية لكل فئة تحدد أكثر الفئات وأقلها مواظبة وأيضًا الفئات متوسطة المواظبة فى العمل.

١٨	٢٤	٣٢	٢٤	١٦
٢٧	١٤	٢٥	٤٠	٣٦
١٢	٢٣	٣٤	٣٨	١٧
١٨	٣٨	٢٤	٤٢	٢٢
٢٧	١٦	٣٥	٢٠	٢٦

جدول رقم (٢-٤)

جدول تكرارى نسبى لأجور العمال

الفئات	التكرار	التكرار النسبى	النسبة المئوية %
١٠-	٢	٢/٢٥ = ٠,٠٨	٨
١٥-	٥	٥/٢٥ = ٠,٢	٢٠
٢٠-	٦	٦/٢٥ = ٠,٢٤	٢٤
٢٥-	٤	٤/٢٥ = ٠,١٦	١٦
٣٠-	٢	٢/٢٥ = ٠,٠٨	٨
٣٥-	٤	٤/٢٥ = ٠,١٦	١٦
٤٠-٤٥	٢	٢/٢٥ = ٠,٠٨	٨
مجموع	٢٥	مجموع نسبى = ١	مجموع مئوى = ١٠٠%

من الجدول يتضح أن أكثر الفئات انتظامًا ومواظبة هى فئتا الأجر (٢٠-)، (١٥-) وأقل الفئات انتظامًا ومواظبة هى فئات الأجر (٤٠-)، (٣٠-)، (١٠-).

جداول التكرار التجمعى Cumulative Frequency Tables:

يستخدم هذا النوع من الجداول التكرارية إذا أراد باحث أن يعرف عدد المفردات التى تقل أو تزيد عن قيمة معينة وذلك نظرًا لأن الجدول التكرارى المتجمع يسهل معرفة تكرارات كل فئة على حدة. ولعل من أهم ما يتميز به الجدول التكرارى التجمعى هو السهولة فى الوصول إلى التكرارات المطلوبة.

وهناك نوعان من التكرارات التجمعية وفقًا لترتيبها إما تصاعديًا فيطلق عليها الجدول التكرارى المتجمع الصاعد، وإما تنازليًا ويطلق عليها الجدول التكرارى المتجمع الهابط.

فلكى يحصل الباحث على قيم بعض المفردات التى تزيد فى القيمة عن قيمة معينة، يقوم بجمع القيم الأقل من الحد الأعلى للفئة المعينة. ومن ثم يقع مجموع تلك التكرارات فى مدى نفس الفئة. ويراعى أن تكون عملية جمع التكرارات الأقل متوالية فى ترتيب يبدأ من أحد طرفى الجدول حتى طرفه الآخر من أجل الحصول على المجموع الكلى للتكرارات.

ففـــى حالـــة التكرار المتجمع الصاعد يكون التجميع من أعلى إلى أسفل ويبدأ الباحـــث بشـــكل تسلســـلى ابتداءً من الفئات الأقل حتى يصل إلى تكرارات الفئات العلـــيا، ويلاحـــظ أن تكون التكرارات المتجمعة فى زيادة مستمرة حتى تصل إلى أكبر تكرار والذى يمثل الحد الأخير فى التسلسل. ويطلق على عمود ترتيب الفئات الصاعدة إما (أقل من الحد الأعلى) أو (الحدود العليا للفئات) وكلاهما صحيح.

وفـــى حالـــة الـــتكرار المتجمع الهابط، يكون التكرار المتجمع للفئة الأخيرة مســـاويًا فى القيمة لتكرارها بالجدول. ويبدأ الباحث فى تجميع التكرارات من أسفل إلى أعلى مبتدئًا بالفئات الكبيرة تنازليًا حتى ينتهى بالفئات الأقل فالأقل، حتى تكون أقـــل قـــيمة تكـــرارية هـــى الفئة الأخيرة بالجدول. وفى كلا النوعين من الجدول الـــتجمعى، يجمع الباحث تكرار كل فئة على مجموع التكرارات السابقة عليها مع رصـــد المجموع الكلى أمام تلك الفئة ويطلق على عمود ترتيب الفئات تنازليًا (الحد الأدنى فأكثر) أو (الحدود السفلى للفئات) وكلاهما صحيح أيضًا.

مثال :

وضح نوعى الترتيب التجمعى لعدد الأفراد القاطنين فى إحدى المبانى السكنية وفقًا للفئات العمرية.

جدول رقم (٢-٥)

جدول تكرارى متجمع صاعد وهابط

الجدول التكرارى		**الجدول التكرارى المتجمع الصاعد**		**الجدول التكرارى المتجمع الهابط**	
فئات (عمرية)	**عدد الأفراد القاطنين**	**الحدود العليا للفئات**	**التكرار المتجمع الصاعد**	**الحدود الدنيا للفئات**	**التكرار المتجمع الهابط**
٢٠-	١٤	أقل من ٢٥	١٤	٢٠ فأكثر	٧٠
٢٥-	٨	أقل من ٣٠	٢٢	٢٥ فأكثر	٥٦
٣٠-	٧	أقل من ٣٥	٢٩	٣٠ فأكثر	٤٨
٣٥-	٥	أقل من ٤٠	٣٤	٣٥ فأكثر	٤١
٤٠-	١٦	أقل من ٤٥	٥٠	٤٠ فأكثر	٣٦
٤٥-٥٠	٢٠	أقل من ٥٠	٧٠	٤٥ فأكثر	٢٠
مجـــ	٧٠				

الجداول المزدوجة :

تشـــتمل الجـــداول المزدوجة على تبويب البيانات وفقًا لمتغيرين بحيث تشمل الصـــفوف تكـــرارات المتغير الأول، بينما تشمل الأعمدة تكرارات المتغير الثانى. مـــثال ذلك الجدول المزدوج التالى الذى يشتمل على حصر عدد الوفيات خلال عام فـــى الـــريف والحضـــر داخل محافظات الوجهين البحرى والقبلى فى المحافظات الحضرية كما يتضح من جدول رقم (٢-٦).

جدول رقم (٢-٦)
توزيع عدد الوفيات (بالألف) على مستوى الجمهورية (ريف وحضر)

الوفيات/ المحافظات	عدد الوفيات (بالألف)		الجملة
	حضر	ريف	
الوجه البحرى	٤٨	٥٢	١٠٠
الوجه القبلى	٥٠	٥٦	١٠٦
المحافظات الحضرية	٥٨	–	٥٨
المجموع	١٥٦	١٠٨	٢٦٤

بالإضـــافة إلـــى مـــا سبق هناك من أنواع الجداول المزدوجة ما يشتمل على ظاهرتيـــن تتصفان بأكثر من نوعين ويراد معرفة العلاقة بينهما. ويطلق على هذا النوع الجداول التوافقية والمثال التالى يوضح جدول توافقى Contingency Table وهـــو يوضـــح تقديرات النجاح لعدد ١٠٠ طالب وطالبة فى مادتى الإحصاء وعلم الاجتماع الصناعى فى قسم الاجتماع:

جدول رقم (٢-٧)
جدول توافقى

مادة الإحصاء مادة الاجتماع الصناعى	ضعيف	مقبول	جيد	ممتاز	المجموع
ضعيف	١١	٧	٥		٢٣
مقبول	٨	٣٠	١٠		٤٨
جيد	٢	١٤	١٠	١	٢٧
ممتاز	–	–	١	١	٢
المجموع	٢١	٥١	٢٦	٢	١٠٠

الجداول التكرارية للبيانات (الكيفية)

تشـــير البـــيانات الكيفية إلى صفات نوعية توضح عناصر الظاهرة موضوع الدراسة مثل الديانة، والمهنة، ومنطقة السكن، والحالة الزواجية والجنسية.

أ – الجداول التكرارية البسيطة:

مثال:

حصــــل أحد الباحثين على البيانات التالية والتى تتعلق بالجنسية لمجموعة من طـــلاب جامعـــة عين شمس. والمطلوب عمل جدول تكرارى بسيط يوضح توزيع الطلاب حسب الجنسية.

مصرى	أجنبى	عربى	مصرى	أجنبى	عربى
مصرى	أجنبى	عربى	مصرى	عربى	مصرى
عربى	عربى	أجنبى	مصرى	أجنبى	مصرى
مصرى	مصرى	أجنبى	مصرى	مصرى	عربى
مصرى	مصرى	عربى	عربى	عربى	أجنبى
مصرى	مصرى	مصرى	عربى	أجنبى	مصرى

خطوات الحل:

١– لتنظيم هذه البيانات فى جدول توزيع تكرارى يلزم عمل جدول تفريغ (جدول رقم ٢–٨) يشتمل على ثلاثة أعمدة كما سبق وأوضحنا.

أ – يتضمن العمود الأول خواص المتغير وهى مصرى، عربى، وأجنبى. ويكون عنوان هذا العمود الجنسية.

ب– يشـــتمل العمـــود الثانى على العلامات حيث يتم تسجيل المشاهدات على شكل خطوط رأسية مائلة تساوى فى عددها التكرارى للصفة الواحدة. فإذا تكررت الصـــفة مرة واحدة يسجل خط رأسى مائل (/)، وإذا ما وصلت عدد العلامات إلـــى أربـــع تكتب أربعة خطوط رأسية مائلة (////)، ثم يكتب الخط الخامس أفقـــيًا لـــيقطع الأربعـــة خطوط المائلة ويعرف هذا الشكل بالحزمة (~~////~~) وعددهـــا خمســـة تكـــرارات. وهكذا يتم تسجيل جميع المشاهدات فى أشكال خطوط رأسية مائلة حتى أربعة تكرارات أو حزمة وتكراراتها.

جـ– أمـــا العمـــود الثالث فيكون بعنوان التكرارات ويرصد فيها عدد تكرارات كل خاصية من خواص المتغير.

جدول رقم (٢-٨)

تفريغ البيانات المتعلقة بجنسية الطلاب

الجنسية	العلامات	التكرارات
مصرى	~~////~~ ~~////~~ ~~////~~ //	١٧
عربى	~~////~~ ~~////~~ /	١١
أجنبى	~~////~~ ///	٨
المجموع		٣٦

٢- مـن الجدول السابق يتم عمل جدول التوزيع التكرارى ويشتمل على عمودين: أولهمـا بعنوان اسم المتغير (الجنسية)؛ وثانيهما يتضمن التكرارات المناظرة لكـل صـفة من صفات المتغير. أى أن جدول التوزيع التكرارى يشتمل فقط علـى العموديـن الأول والثالـث من جدول تفريغ البيانات الخام. وذلك على النحو التالى :

جدول رقم (٢-٩)

التوزيع التكرارى للطلاب حسب الجنسية

الجنسية	ك	%
مصرى	١٧	٤٧,٢
عربى	١١	٣٠,٦
أجنبى	٨	٢٢,٢
المجموع	٣٦	١٠٠

ب- الجداول التكرارية المزدوجة للبيانات الكيفية:

تشـتمل الجداول التكرارية المزدوجة على متغيرين كالتوزيع التكرارى لأفراد العيـنة حسـب الأصول الريفية الحضرية والجنسية، أو حسب النوع والمهنة، أو الجنسية والنوع.

مثال :

عـند دراسة الجنسية والحالة الزواجية لعينة عشوائية تم سحبها من إحدى الشـركات الاسـتثمارية بمدينة العاشر من رمضان، كانت النتائج على النحو التالى:

مصريون			عرب			أجانب		
متزوج	متزوج	لم يسبق	متزوج	لم يسبق	متزوج	متزوج	لم يسبق	مطلق
متزوج	لم يسبق	له الزواج	أرمل	له الزواج	متزوج	أرمل	له الزواج	متزوج
لم يسبق	له الزواج	لم يسبق	مطلق	متزوج	أرمل	متزوج	متزوج	
له الزواج	متزوج	له الزواج	متزوج	أرمل	متزوج	متزوج	متزوج	
لم يسبق	أرمل	أرمل	متزوج	أرمل		متزوج	لم يسبق	
له الزواج	متزوج	مطلق	متزوج	متزوج		أرمل	له الزواج	
أرمل	أرمل	متزوج	لم يسبق	لم يسبق		لم يسبق	متزوج	
مطلق		متزوج	له الزواج	له الزواج		له الزواج	مطلق	

والمطلوب عمل جدول تكرارى مزدوج لعرض هذه البيانات.

خطوات الحل:

١- عمـــل جـــدول تفريغ مزدوج بحيث يشتمل العمود على صفات المتغير الأول الحالة الزواجية (لم يسبق له الزواج، متزوج، أرمل، مطلق). ويشتمل الصف علـــى خـــواص المتغير الثانى الجنسية (مصريون، عرب، وأجانب). وتشتمل كـــل خاصـــية فى الصف على عمودين أولهما للعلامات وثانيهما للتكرارات المـــناظرة لعدد القراءات المسجلة بالعلامات أمام كل خاصية فى العمود كما يتضح من الجدول رقم (٢-١٠).

جدول رقم (٢-١٠)
تفريغ البيانات المتعلقة بالجنسية والحالة الزواجية

الجنسية / الحالة الزواجية	مصريون		عرب		أجانب		المجموع
	العلامات	ك	العلامات	ك	العلامات	ك	ك
لم يسبق له الزواج	~~////~~	٥	///	٣	///	٣	١١
متزوج	~~////~~ //	٧	~~////~~ ////	٩	~~////~~ ///	٨	٢٤
أرمل	////	٤	////	٤	//	٢	١٠
مطلق	//	٢	/	١	//	٢	٥
المجموع		١٨		١٧		١٥	٥٠

٢- عمل جدول توزيع تكرارى لمفردات العينة حسب الجنسية والحالة الزواجية من جدول التفريغ السابق على النحو التالى:

جدول رقم (٢-١٢)
التوزيع التكرارى لمفردات العينة حسب الجنسية والحالة الزواجية

الجنسية / الحالة الزواجية	مصريون	عرب	أجانب	المجموع
لم يسبق له الزواج	٥	٣	٣	١١
متزوج	٧	٩	٨	٢٤
أرمل	٤	٤	٢	١٠
مطلق	٢	١	٢	٥
المجموع	١٨	١٧	١٥	٥٠

جدول رقم (٢-١٢)
التوزيع التكرارى المئوى (النسبى) لأفراد العينة
حسب الجنسية والحالة الزواجية

الجنسية / الحالة الزواجية	مصروين		عرب		أجانب		المجموع	
	ك	%	ك	%	ك	%	ك	%
لم يسبق له الزواج	٥	٢٧,٨	٣	١٧,٦	٣	٢٠,٠	١١	٢٢,٠
متزوج	٧	٣٨,٩	٩	٥٢,٩	٨	٥٣,٣	٢٤	٤٨,٠
أرمل	٤	٢٢,٢	٤	٢٣,٥	٢	١٣,٣	١٠	٢٠,٠
مطلق	٢	١١,١	١	٥,٩	٢	١٣,٣	٥	١٠,٠
المجموع	١٨	٣٦,٠	١٧	٣٤,٠	١٥	٣٠,٠	٥٠	١٠٠

وتحسب النسبة المئوية لكل خلية على النحو الآتى:

$$\text{النسبة المئوية} = \frac{\text{تكرار الخلية}}{\text{المجموع الكلى}} \times ١٠٠$$

ويتم التعليق على الجدول باستخدام النسب المئوية.

ملاحظات عامة :

تعتمد بنية الجدول الإحصائى على الهدف الذى يتم من أجله اختيار نوع الجدول وتبويب البيانات. وفيما يلى بعض الملاحظات التى ينبغى على الدارس أن يأخذها بعين الاعتبار عند استخدام الجداول الاحصائية فى بحثه.

أ - ضرورة كتابة عنوان واضح ومختصر بحيث يوضح بإيجاز ما يتضمنه الجدول من علاقات.

ب- فى حالة اشتمال تقرير البحث على عدد قليل من الجداول الإحصائية يفضل استخدام أرقام مسلسلة يرقم بها الجداول من بداية التقرير حتى نهايته (جدول رقم ١)، (جدول رقم ٢) وهكذا حتى نهاية البحث أما فى حالة اشتمال البحث على عدد كبير من الجداول يفضل ترقيم جداول كل فصل على حدة وفى هذه الحالة يكون للجدول رقمان، يكون الرقم الأول دالاً على رقم الفصل والرقم الثانى دالاً على رقم الجدول فى الفصل (مثال: جدول رقم ٣-٥) يكون دالاً على انه الجدول رقم (٥) فى الفصل الثالث.

جـ- الحرص على تحديد نوع الوحدة القياسية للمعلومات والبيانات التى يتضمنها الجدول مثل وحدات الأطوال (سنتيمتر، متر، كيلو متر وغيرها) أو وحدات الوزن (كيلو جرام، رطل، طن وغيرهم) وكذلك وحدات العمر (سنة أو يوم أو شهر). كما يراعى فى بعض الأحيان ضرورة تحديد اتجاهات القيم بوضع إشارات جبرية سالبة أو موجبة أمام تلك الأرقام. كذلك فى حالة استخدام النسب فمن الضرورى تحديد تلك النسب هل هى مئوية أو فى الألف أو فى المليون وهكذا.

د - يجب أن تكون عناوين الصفوف والأعمدة مختصرة وواضحة.



الفصل الثالث

التمثيل البيانى للبيانات

أولاً: نظام المحاور الإحداثية.

ثانيًا: التمثيل البيانى للبيانات المتقطعة.

١ – الأعمدة:

أ – الأعمدة البسيطة.

ب– الأعمدة المزدوجة.

جـ– الأعمدة المجزأة.

د– الأعمدة المنزلقة.

٢ – الدائرة.

ثالثًا: التمثيل البيانى للبيانات المتصلة.

١ – المدرج التكرارى.

٢ – المضلع التكرارى.

٣– المضلع التكرارى التجمعى.

٤ – المنحنى التكرارى.

٥– المنحنيات المتجمعة.

٦– المنحنى المتجمع الهابط.

٧– المنحنى المتجمع الصاعد.

الفصل الثالث
التمثيل البيانى للبيانات

مقدمة :

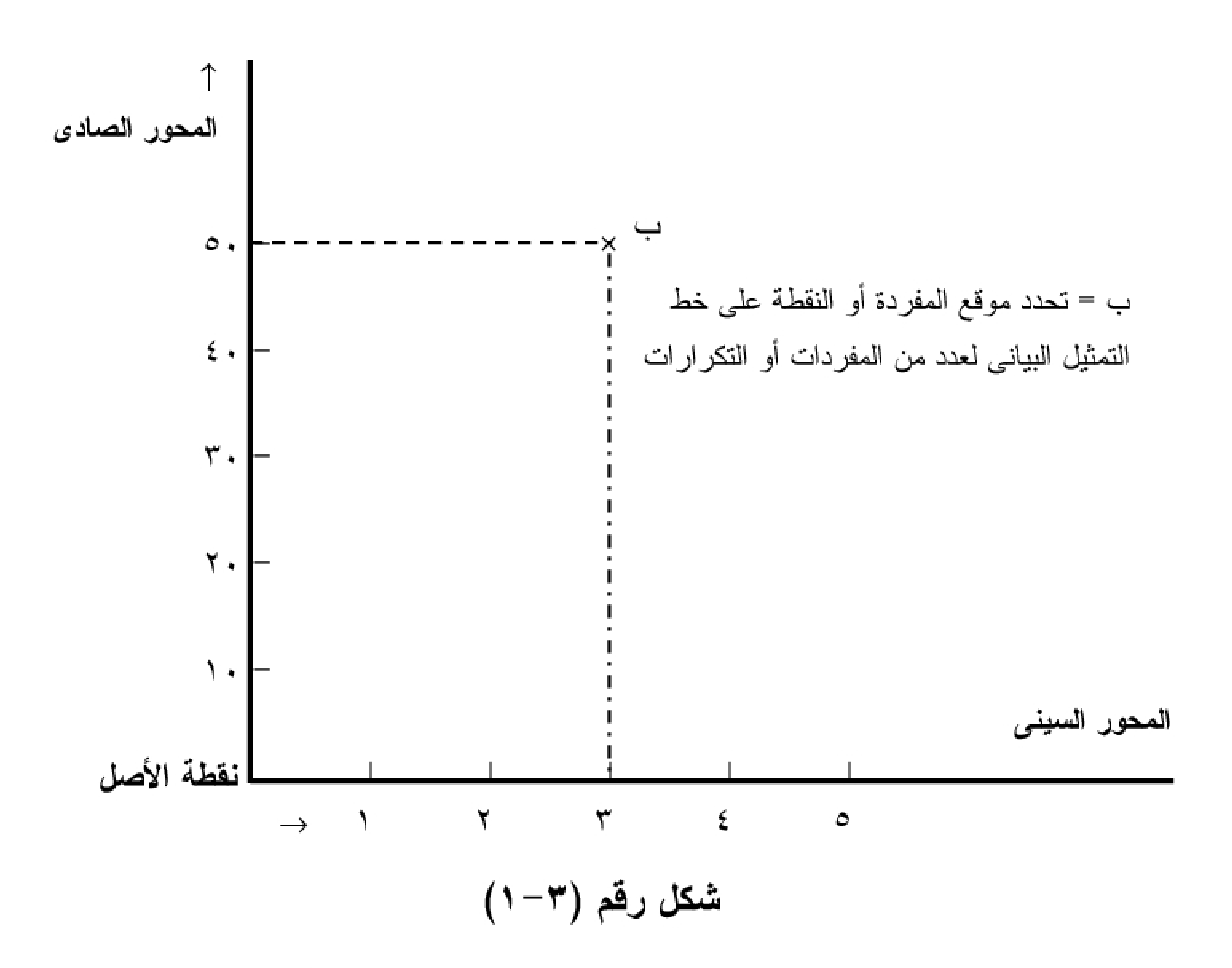
قــبل أن نتــناول الأشكال البيانية المختلفة للتوزيعات التكرارية نرى ضرورة شــرح نظــام المحاور الإحداثية المشتركة بإيجاز، حيث نجد أنه يتضمن محورين أحدهمــا يسمى المحور الأفقى أو المحور السينى، بينما يطلق على المحور الرأسى المحــور الصادى، ومنطقة التقاء المحورين هى نقطة الأصل (المنطقة الصفرية)، وبالإضــافة إلــى نظــام المحاور الإحداثية المشتركة فسوف نتناول بالشرح أيضًا التمثيل البيانى للبيانات المتقطعة والتمثيل البيانى للبيانات المتصلة.

ويهــدف هذا الفصل إلى تعريف الدارس بكيفية رسم الأشكال البيانية للبيانات المــتقطعة والمتمــثلة بــالأعمدة البسيطة، والمتلاصقة، والمجزأة والدائرة. بالنسبة للبــيانات المتصلة فيكون التعبير عنها بيانيًا باستخدام المنحنى التكرارى، والمضلع الــتكرارى، والمــدرج التكرارى. والمنحنى المتجمع الصاعد، والمنحنى المتجمع الهابط.

أولاً : نظام المحاور الاحداثية Cartesian Coordinate System :

يــتكون هذا النظام المشترك من مقياسين رقميين يتعامد كل منهما على الآخر بزاوية قدرها ٩٠ درجة، ويطلق على المقياس الأول المقياس السينى. ويتمثل بخط أفقــى يــتم تقســيمه إلى مسافات متساوية الأبعاد ابتداءً من نقطة الالتقاء بين هذا المقــياس والمقــياس الرأسى المتعامد عليه ويطلق على نقطة الالتقاء نقطة الأصل Origin، كذلك يتم تقسيم المحور الرأسى إلى مسافات متساوية الأبعاد ويطلق عليه المحــور الصادى. معنى نظام المحاور الإحداثية المشتركة أن المفردة الواحدة يتم توقــيعها وفــق قيمتين مناظرتين لها حيث إن لكل قيمة على المقياس السينى (قيمة مســتقلة)، وقــيمة أخرى مناظرة تابعة على المحور الصادى الرأسى (قيمة تابعة) وطــبقًا للارتــباط بيــن القيمتيــن للمفردات يتحدد شكل المنحنى أو الرسم البيانى ويحصــل على علاقة إما خطية أو انحنائية بين المتغيرين (س، ص). حيث (س) تســاوى طول المسافة ابتداءً من نقطة الأصل حتى قيمتها المعطاة. و(ص) تساوى طــول المســافة ابــتداءً من نقطة الأصل حتى قيمتها المعطاة كما يتضح ذلك من الشــكل رقم (٣-١) بالنسبة لمفردة واحدة تحددها قيمتان (٣، ٥٠) وطريقة تحديد

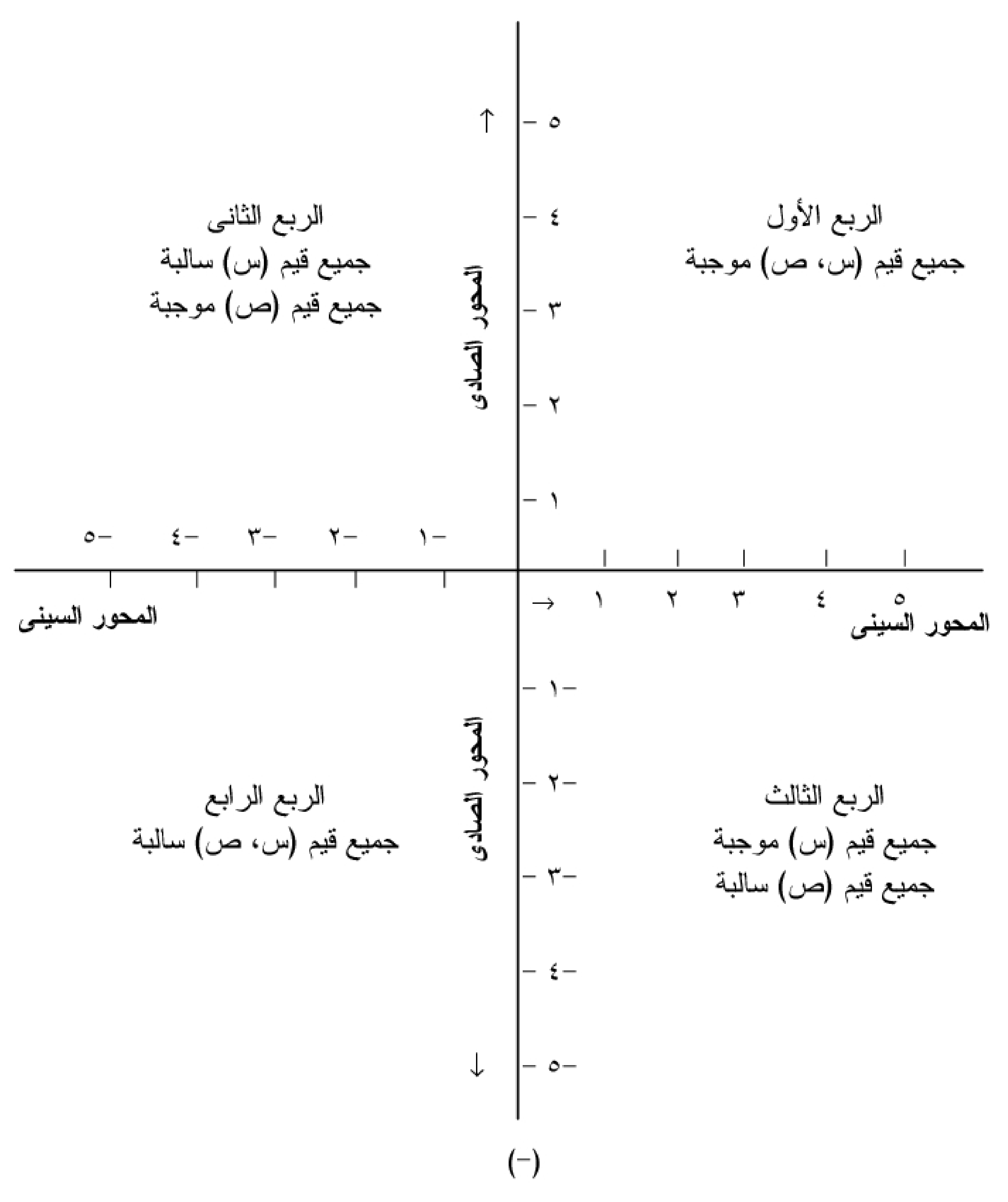
موضـــع المفردة بإحداثييها (٣، ٥٠)، أن نبدأ فى قياس ثلاث وحدات على المحور الســـينى بدايـــة مـــن نقطة اليسار وعند القيمة (٣) يقام خط رأسى موازى للمحور الصـــادى. ثـــم بعـــد ذلك نقوم بقياس القيمة (ص) وهى (٥٠) لنفس المفردة على المحـــور الصادى وعند تلك القيمة نرسم خطًا أفقيًا موازيًا للمحور السينى، فيلتقى الخطـــان فـــى نقطة (ب) التى تحدد موضع المفردة المعطاة كما فى الشكل التالى، وفـــى حالة تعدد المفردات أو التكرارات نكرر نفس العمل لكل مفردة فنحصل فى النهاية على الشكل البيانى أو التمثيل البيانى للتوزيعات التكرارية.



شكل رقم (٣-١)

وإذا كانـــت قيمـــتا المفـــردة وهمـــا (س، ص) سالبتين أو موجبتين أو كانت إحداهمـــا سالبة والأخرى موجبة؛ فإن المحاور الإحداثية تقسم إلى أربعة مساحات، اثنين أعلى المحور الأفقى بحيث تكون الأولى ناحية اليمين من نقطة الأصل جميع قـــيمها موجـــبة والثانـــية إلـــى اليسار من نقطة الأصل ولها نفس المقياس بأبعاده المتســـاوية وتكون جميع القيم الواقعة بداخلها سالبة للقيمتين (س، ص) لكل مفردة كذلك توجد مساحتان أسفل الخط الأفقى يقسمهما امتداد المحور الصادى متجهًا إلى الأســـفل مـــن نقطـــة الأصل بحيث تكون جميع قيم (س) موجبة وجميع قيم (ص)

ســـالبة للمفردات، وهذه المساحة تقع على يمين المحور الصادى المتجه لأسفل. أما المســـاحة الـــرابعة فـــتقع أســـفل المحور الأفقى، وعلى اليسار من امتداد المحور الصـــادى المـــتجه إلـــى أســـفل كما يتضح ذلك من الشكل رقم (٣-٢). وفى هذه المساحة الرابعة تكون جميع قيم (س) وجميع قيم (ص) سالبة.



شكل رقم (٣-٢) محاور الإحداثيات فى مستويين (س، ص)

ثانياً: التمثيل البيانى للبيانات المتقطعة Discrete data:

يمكــن التعبــير عــن المتغيرات المتقطعة إما بقياسات اسمية Nominal أو بقياســات رتبية Ordinal ففى حالة القياس الاسمى، يتم تقسيم المتغير إلى خواصه أى فــئات نوعية، كلٌ منها مستقل بذاته، ويضم الأفراد المتماثلين فى هذه الخاصية مثال: الحالة الزواجية فإنها تشتمل على الفئات التالية :

١- أعزب.
٢- متزوج.
٣- مطلق.
٤- أرمل.

وفــى حالة استخدام المقياس الرتبى يكون ترتيب المجموعات وفقًا للتميز فى الصــفة بحيث نقول على سبيل المثال إن المجموعة رقم (أ) أكبر من (أو أقل من) المجموعــة رقــم (ب). أيضًــا يمكــن اســتخدام الأشكال البيانية التالية فى تمثيل التوزيعات التكرارية المتقطعة.

١ - الأعمدة البسيطة:

تســتخدم لتمثــيل قــيم المشاهدات لظاهرة واحدة. قد تكون المشاهدات مقيسة بالنسبة للزمن أو غير ذلك.

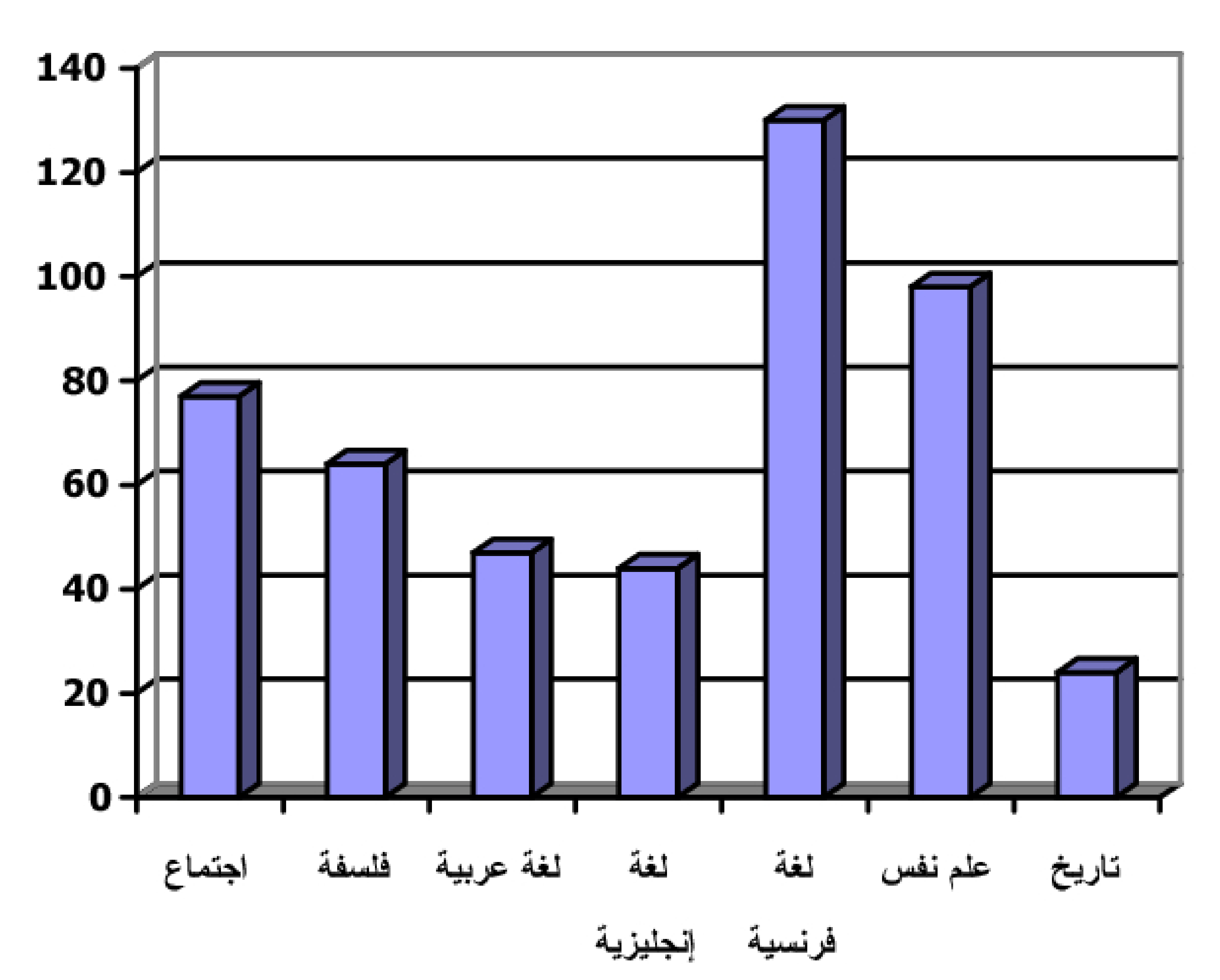
مثال:

الجــدول الآتــى يتضمن عدد الطالبات فى الأقسام المختلفة فى إحدى الكليات للفــرقة الرابعة خلال العام الجامعى ٢٠٠٨/٢٠٠٩. والمطلوب تمثيل هذه البيانات باستخدام الأعمدة البسيطة.

جدول رقم (٣-١)

توزيع طالبات الفرقة الرابعة حسب التخصص عام ٢٠٠٨/٢٠٠٩

التخصص	اجتماع	فلسفة	لغة عربية	لغة إنجليزية	لغة فرنسية	علم نفس	تاريخ
عدد الطالبات	٧٧	٦٤	٤٧	٤٤	١٣٠	٩٨	٢٤



شكل يوضح توزيع طالبات الفرقة الرابعة حسب التخصص عام ٢٠٠٨/٢٠٠٩

٢- الأعمدة المزدوجة (المتلاصقة):

تستخدم الأعمدة المزدوجة فى حالة المقارنة بين ظاهرتين أو أكثر كالمقارنة بين الدخل والاستهلاك، وعدد الطلبة وعدد الطالبات فى المدرسة أو الجامعة.

تُمثل الأعمدة المتلاصقة بأن يمثل كل ظاهرة عمود يلاصق عمود أو مستطيل الظاهرة الثانية. ويرسم كل مستطيل بلون مختلف أو يظلل بخطوط مائلة أو لون داكن عن الآخر الذى قد يكون غير مظلل.

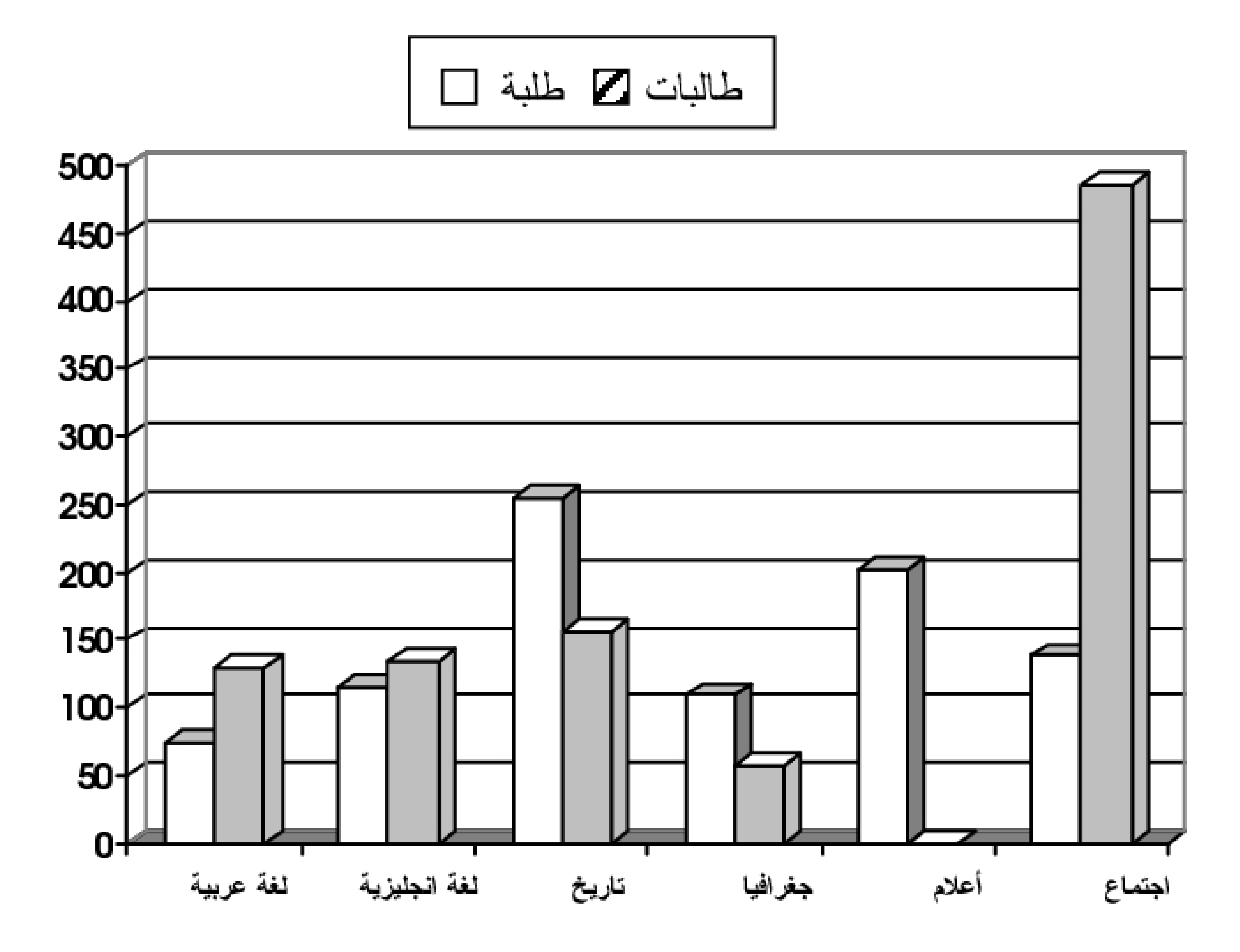
مثال:

استخدم الأعمدة المتلاصقة فى تمثيل البيانات الواردة فى الجدول رقم (٣-٢) لعدد الطلبة والطالبات حسب التخصص فى إحدى الكليات المصرية خلال العام الجامعى ٢٠٠٨/٢٠٠٩.

جدول رقم (٣-٢)

توزيع طلبة وطالبات الفرقة الرابعة حسب التخصص فى إحدى الكليات ٢٠٠٨/٢٠٠٩

التخصص	لغة عربية	لغة إنجليزية	تاريخ	جغرافيا	إعلام	اجتماع
طلبة	٧٤	١١٥	٢٥٤	١٠٩	٢٠١	١٣٨
طالبات	١٢٩	١٣٤	١٥٥	٥٧	-	٤٨٥



شكل يوضح أعداد الطلاب والطالبات حسب التخصص فى إحدى الكليات المصرية حسب التخصص خلال العام الجامعى ٢٠٠٨/٢٠٠٩

٣- الأعمدة المجزأة:

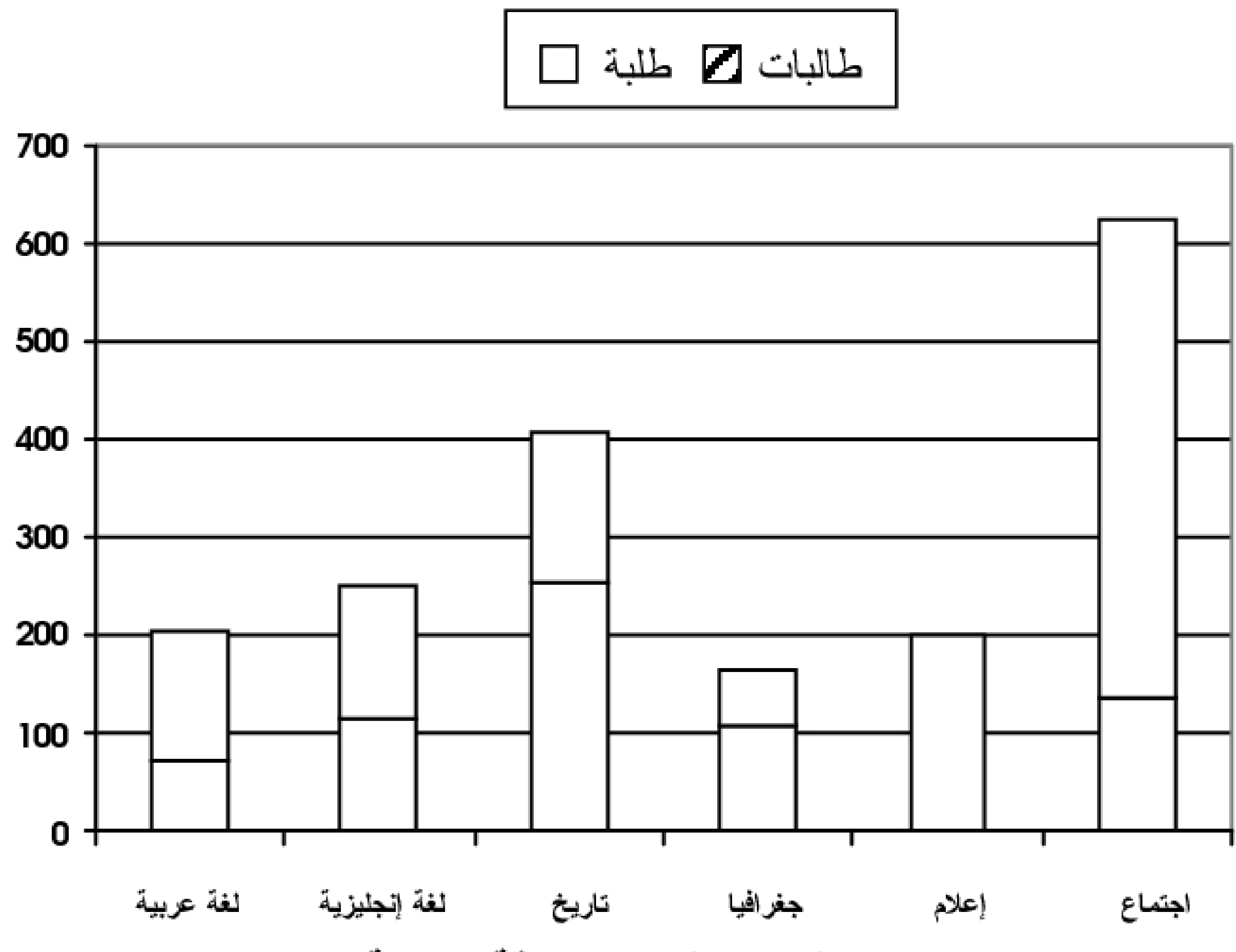
تستخدم الأعمدة المجزأة فى حالة وجود أكثر من ظاهرة مثلما تستخدم الأعمدة المتلاصقة. لكن فى حالة الأعمدة المجزأة يستخدم عمود واحد لمجموع القيم لبيانات الظاهرتين. ثم يُقسَّم المستطيل بنسبة عدد البيانات لكل ظاهرة.

مثال:

استخدم الأعمدة المجزأة فى تمثيل البيانات الواردة فى المثال السابق لعدد الطلاب حسب التخصص.

الحل :يعاد الجدول بعد إضافة خانة للمجموع (البنون + البنات) فى كل تخصص على حدة لمثل طول العمود أو المستطيل.

التخصص	لغة عربية	لغة إنجليزية	تاريخ	جغرافيا	إعلام	اجتماع
طلبة	٧٤	١١٥	٢٥٤	١٠٩	٢٠١	١٣٨
طالبات	١٢٩	١٣٤	١٥٥	٥٧	–	٤٨٥
المجموع	٢٠٣	٢٤٩	٤٠٩	١٦٦	٢٠١	٦٢٣



شكل يوضح توزيع طلاب وطالبات الفرقة الرابعة حسب التخصص

٤- استخدام الدائرة فى العرض البيانى للبيانات:

مثال :

يوضح الجدول الآتى عدد المبعوثين للدراسة خارج مصر فى تخصصات مختلفة.

المطلوب: تمثيل البيانات باستخدام الدائرة.

التخصص	مهندسون	أطباء	مدرسون	محاسبون	المجموع
العدد	٢٢٠	٣٥٠	٢٠٠	١٣٠	٩٠٠

خطوات الحل:

- حساب قطر الدائرة = ق = $\sqrt{\text{المجموع الكلى}}$.

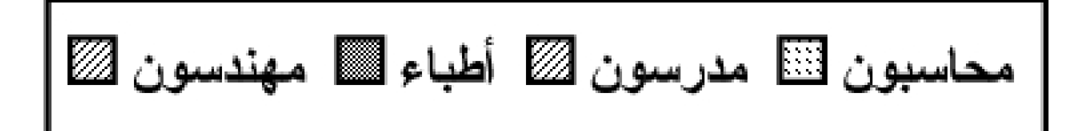
∴ ق = $\sqrt{٩٠٠}$ = ٣٠

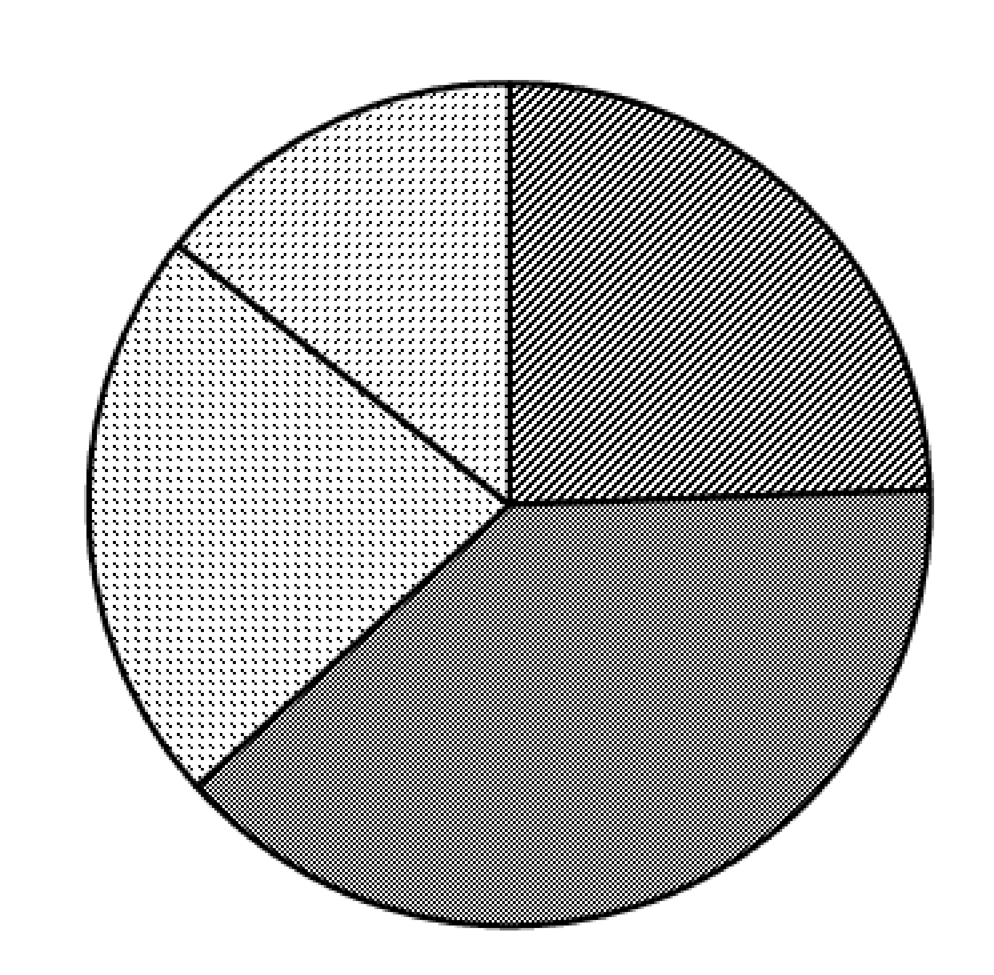
- حساب نصف القطر = $\frac{ق}{٢} = \frac{٣٠}{٢} = ١٥$

- يــراعى أخــذ مقــياس رسم مناسب حتى تكون الدائرة معقولة المسافة على الورقة فمقياس الرسم المناسب هنا هو ٥ : ١ بمعنى أن كل خمسة سنتيمرات يقابلها سنتيمتر واحد.

∴ نق (للرسم) = $\frac{١٥}{٥}$ = ٣ سم

- نبدأ فى الرسم كما يلى :
- نقوم بتحديد الزوايا المركزية وذلك على النحو الآتى:





القطاعات الدائرة :

قطاع المهندسين = $\frac{٢٢٠}{٩٠٠}$ × ٣٦٠° = ٨٨°

قطاع الأطبــاء = $\frac{٣٥٠}{٩٠٠}$ × ٣٦٠° = ١٤٠°

قطاع المدرسين = $\frac{٢٠٠}{٩٠٠}$ × ٣٦٠° = ٨٠°

قطاع المحاسبين = $\frac{١٣٠}{٩٠٠}$ × ٣٦٠° = ٥٢°

ملحوظ : لابد أن يكون مجموع الزوايا = ٣٦٠°

وفى حالة التكرار المئوى تحول النسب المئوية إلى الدرجات (درجات الدائرة) على النحو التالى:

التخصص	ك	%	تحويل النسب إلى درجة الدائرة
مهندسون	٢٢٠	٢٤,٤	٢٤,٤ × ٣,٦ = ٨٧,٨٤
أطباء	٣٥٠	٣٨,٩	٣٨,٩ × ٣,٦ = ١٤٠,٠٤
مدرسون	٢٠٠	٢٢,٢	٢٢,٢ × ٣,٦ = ٧٩,٩٢
محاسبون	١٣٠	١٤,٤	١٤,٤ × ٣,٦ = ٥١,٨٤
مجـــ		١٠٠	٣٦٠ تقريبًا

ثالثاً: التمثيل البيانى للبيانات المتصلة :

١– المدرج التكرارى :

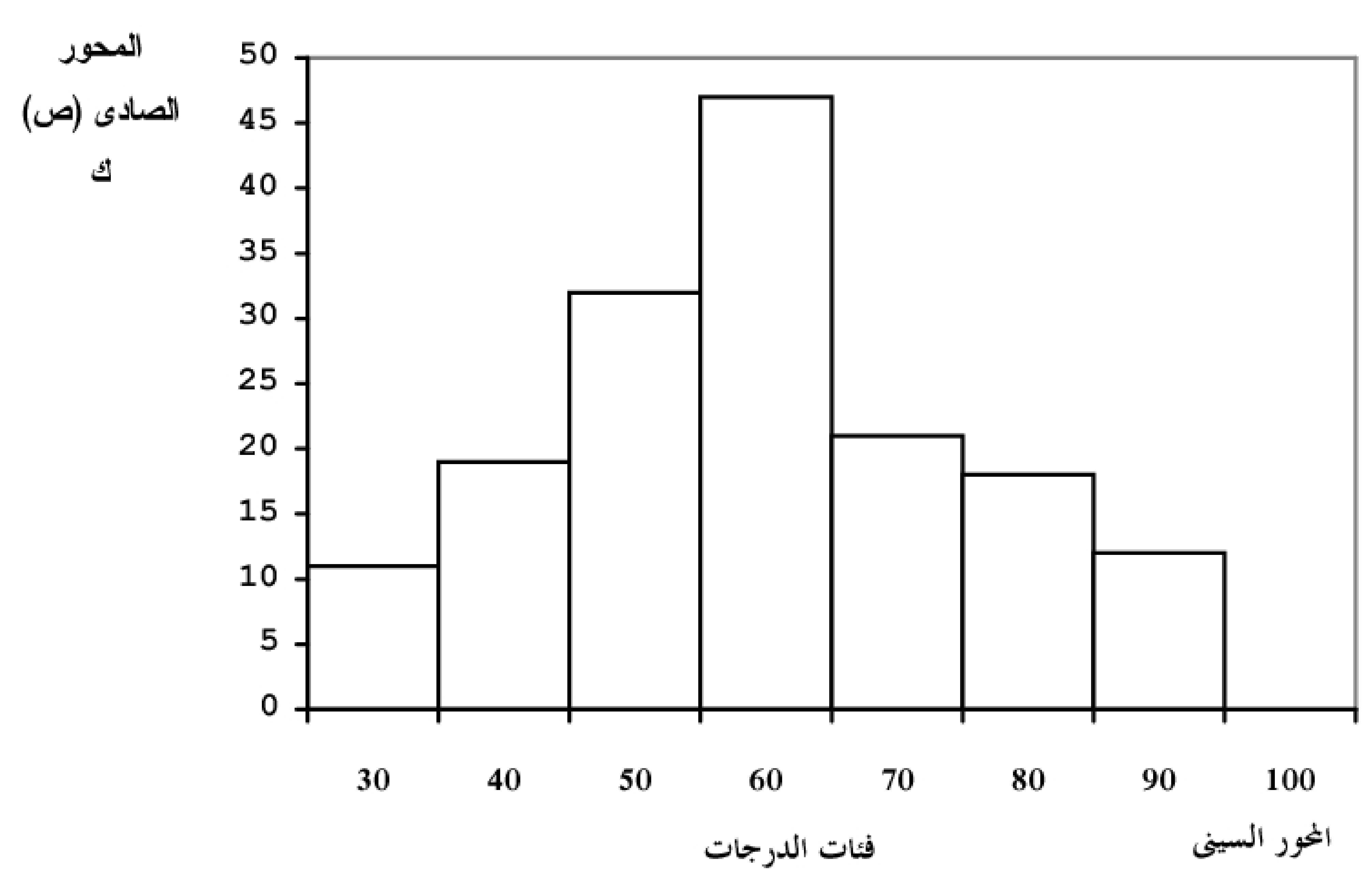
مثال :

فيما يلى توزيع تكرارى لدرجات امتحان مادة مبادئ الإحصاء لطالبات السنة الأولى قسم الاجتماع.

المطلوب: عرض التوزيع بيانيًا باستخدام الطرق الملائمة.

فئات الدرجات	عدد الطالبات (ك)
٣٠–	١١
٤٠–	١٩
٥٠–	٣٢
٦٠–	٤٧
٧٠–	٢١
٨٠–	١٨
٩٠–١٠٠	١٢
مجـــ	١٦٠

أمـــا عن كيفية رسم المدرج التكرارى، فيتم باستخدام محاور الاحداثيات (س، ص) ونختار المحور السينى للفئات والمحور الصادى للتكرارات. ويتكون المدرج التكرارى بواسطة رسم مجموعة من المستطيلات تكون طول قاعدتها على المحور الســـينى مساويًا لطول الفئة وفى هذا المثال نجد أن طول الفئة ثابت ومقداره ١٠. أمـــا ارتفاع المستطيل فقيمته هى التكرار المناظر لكل فئة، ومن ثم نجد أن مساحة كـــل مســـتطيل تتناســـب طرديًا مع التكرار المناظر له. وبالتالى نحصل على عدة مســـتطيلات متجاورة ويطلق على هذا الشكل المدرج التكرارى كما يتضح ذلك من الشكل رقم (٣-٣)



شكل رقم (٣-٣) المدرج التكرارى لدرجات امتحان مبادئ الإحصاء

٢- المضلع التكرارى Frequency polygon :

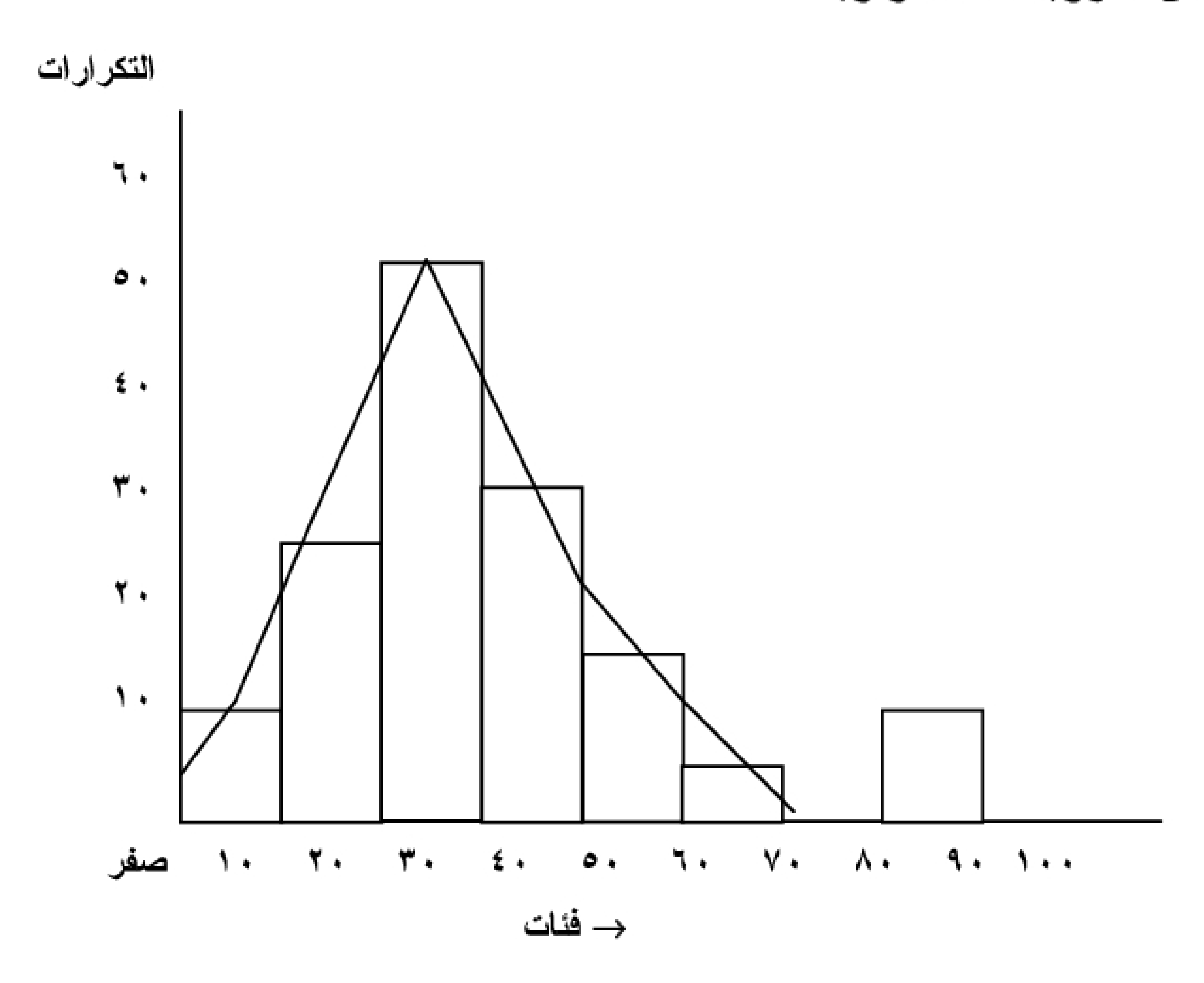
يعتبر المضلع التكرارى الأسلوب الثانى لتمثيل التوزيعات التكرارية المتصلة بيانيًا، حيث يشابه إلى حد بعيد المدرج التكرارى كما قد يستنبط منه. وبدلاً من أن يعـــتمد المـــدرج الـــتكرارى على توزيعات تكرارية مقفلة حيث يتحدد حدى الفئة (الأدنى والأعلى) ويمثلان ضلعى المستطيل، فإن رسم المضلع التكرارى يقوم على فكـــرة استخدام مركز الفئة. ومن ثم يعتمد المضلع التكرارى على نقطة فى مركز

الفــئة وهــذه النقطة إما أن تتحدد فى المدرج التكرارى بتنصيف للخط العلوى لكل مستطيل والموازى للمحور الأفقى فى نقطة تمثل مركز الفئة لكل تكرار مقابل على المحــور الصادى. أو أنه يمكن حسابها بقسمة مجموع قيمتى حدى الفئة (الأدنى + الأعلــى) وقســمة الــناتج على (٢) وذلك فى حالة عدم وجود المدرج التكرارى. وبتوصــيل النقط المتوسطة بخطوط مستقيمة (باستخدام المسطرة) بين كل نقطتين متتاليتين نحصل على شكل المضلع التكرارى.

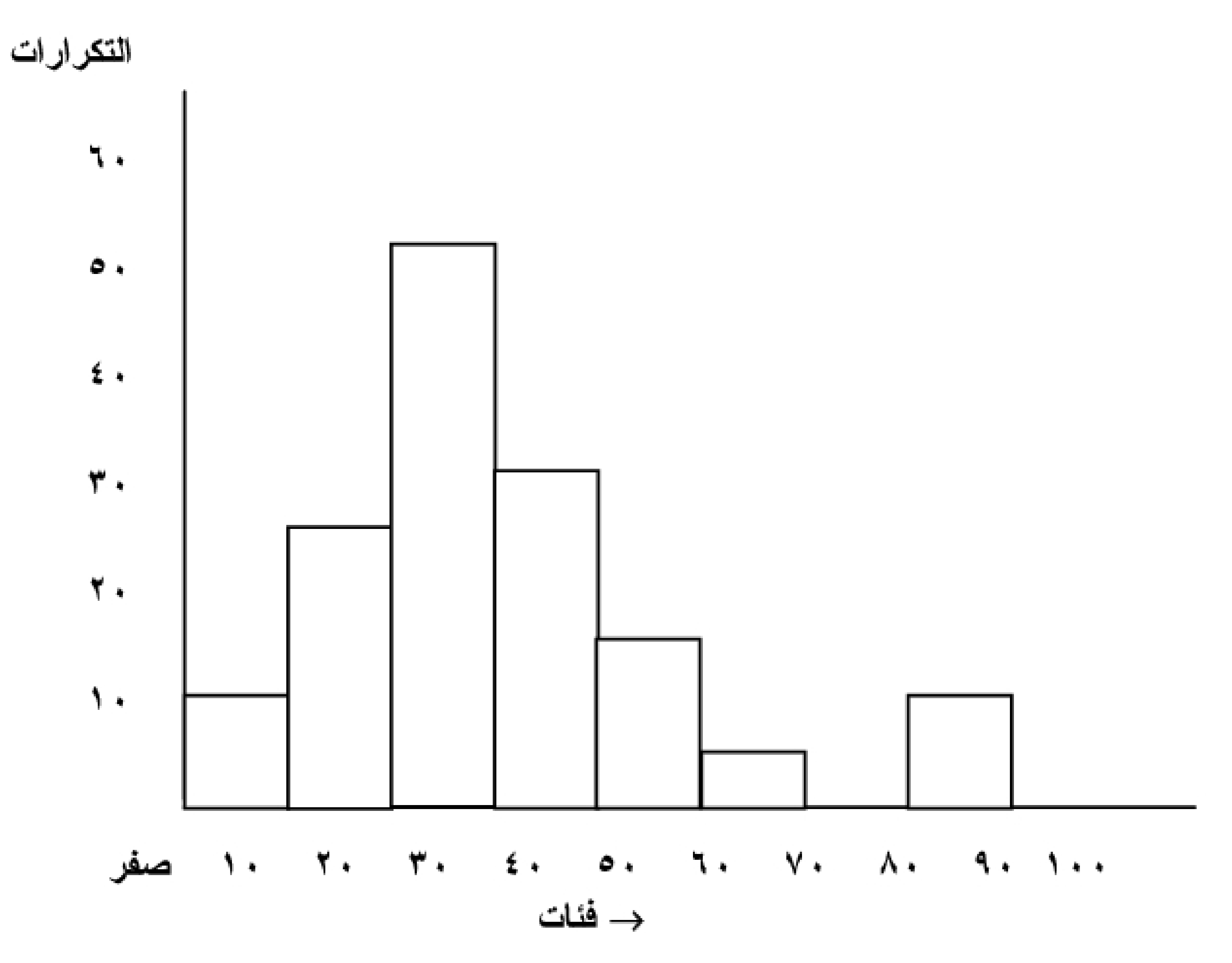
أ - رسم المضلع التكرارى من المدرج التكرارى:

الشــكل رقــم (٣-٤) يبيــن المدرج التكرارى لتوزيع تكرارات متصلة ذات أطوال فئات متساوية.

الشــكل رقم (٣-٥) يبين المضلع التكرارى الناتج من تنصيف الضلع الأفقى والعلــوى لكل مستطيل والتوصيل بخطوط مستقيمة بين كل نقطتين متتاليتين منهما لنفس التوزيعات التكرارية.



شكل رقم (٣-٤) رسم المضلع التكرارى من المدرج التكرارى



شكل رقم (٣-٥) تمثيل التكرارات بالمدرج التكرارى

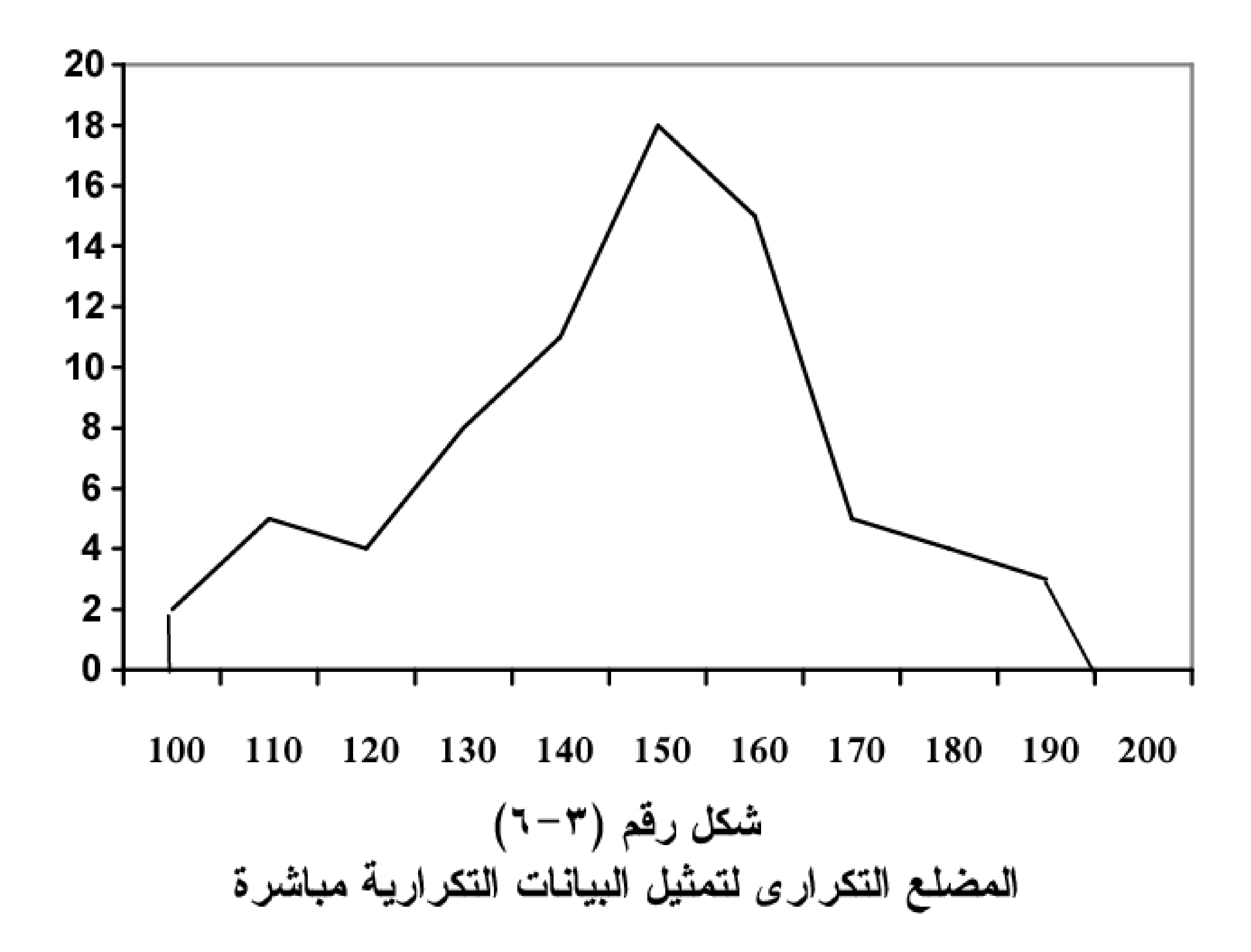
ب- رسم المضلع التكرارى من التوزيعات التكرارية مباشرة :

مثال:

ارسم المضلع التكرارى للتوزيعات التكرارية لأوزان عينة من الأفراد (الأوزان بالرطل).

ف	ك
١٠٠-	٢
١١٠-	٥
١٢٠-	٤
١٣٠-	٨
١٤٠-	١١
١٥٠-	١٨
١٦٠-	١٥
١٧٠-	٥
١٨٠-	٤
١٩٠-٢٠٠	٣
المجموع	٧٥

الحل: يــتم استخدام مراكز الفئات وتوضع النقاط أمام التكرارات المناظرة لكل فئة كما يتضح من الشكل رقم (٣-٦)



شكل رقم (٣-٦)
المضلع التكرارى لتمثيل البيانات التكرارية مباشرة

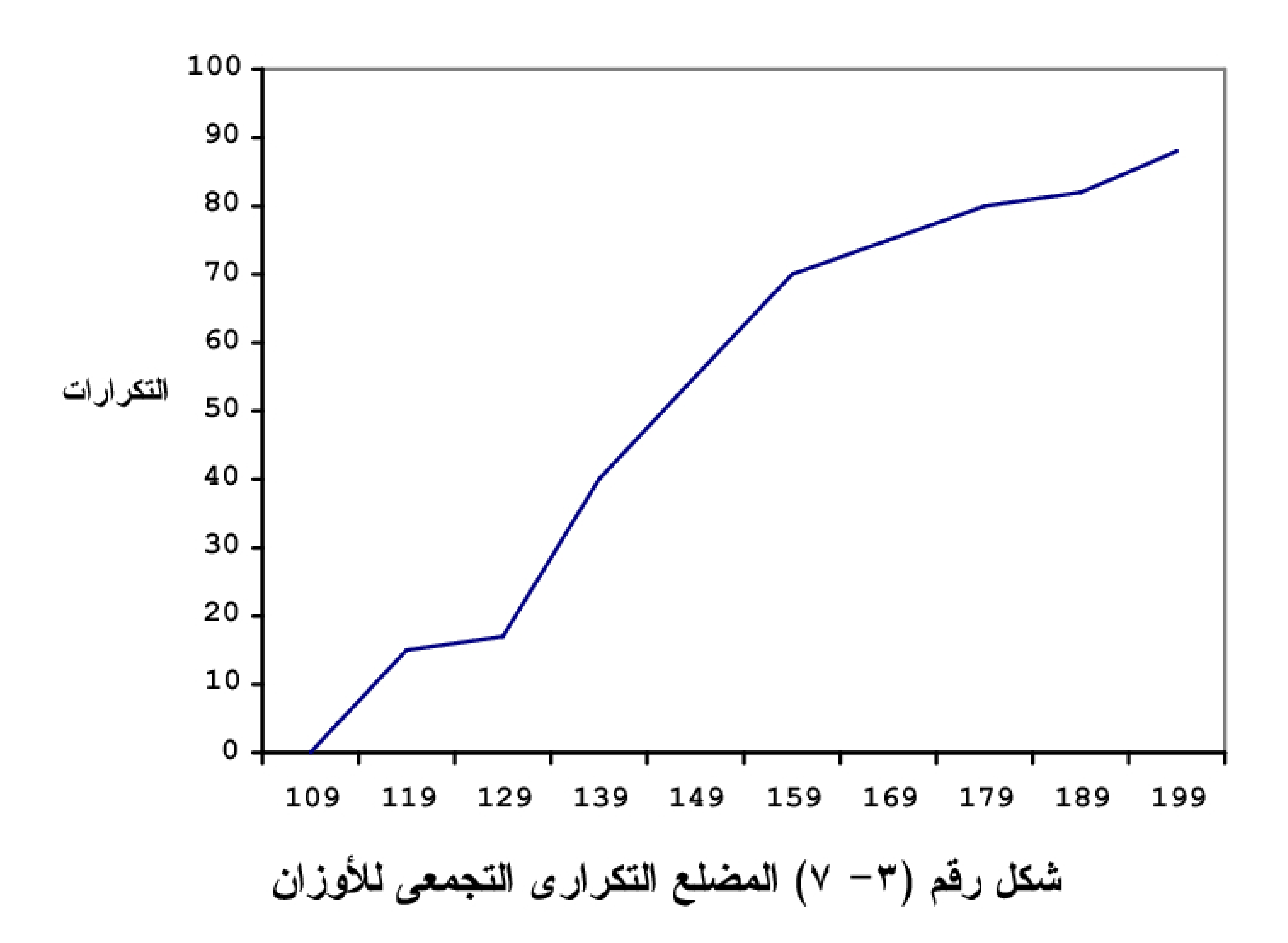
ويفضل دائمًا إقفال المضلع التكرارى، بمعنى أن نهايتيه تلتقيان مع المحور السينى، ويمكــن ذلك بأن يفترض الدارس وجود فئتين إضافيتين على الفئات الأصلية التى يشتمل عليها المدى، شريطة أن يضع أول فئة منهما سابقة لأول فئة أصلية، والفئة الثانــية يضـــعها بعد آخر فئة فى المدى. ويمكن للدارس بالإضافة إلى ما سبق أن يفــترض قيمة تساوى صفرًا لتكرارات كل فئة منهما ويشترط أن تتساويا فى طول الفئة مع باقى الفئات الأصلية.

بعض استخدامات المضلع التكرارى:

المضلع التكرارى التجمعى Cumulative Frequency Polygon:

يمكــن ترتيب التوزيعات التكرارية تجمعيًا تصاعديًا أو تنازليًا، ورسم المضلع التكرارى التجمعى. ويعد هذا الشكل أحد استخدامات المضلع التكرارى، ففى المثال الســابق للــتوزيعات الــتكرارية لــلأوزان (بالرطل) يمكن التمثيل البيانى بمضلع تكرارى تجمعى صاعد (من أقل قيمة إلى أعلى قيمة). إلا أن طريقة رسم المضلع الــتكرارى الــتجمعى لا تعــتمد على قيم النقط المتوسطة بل تعتمد على قيم الحد

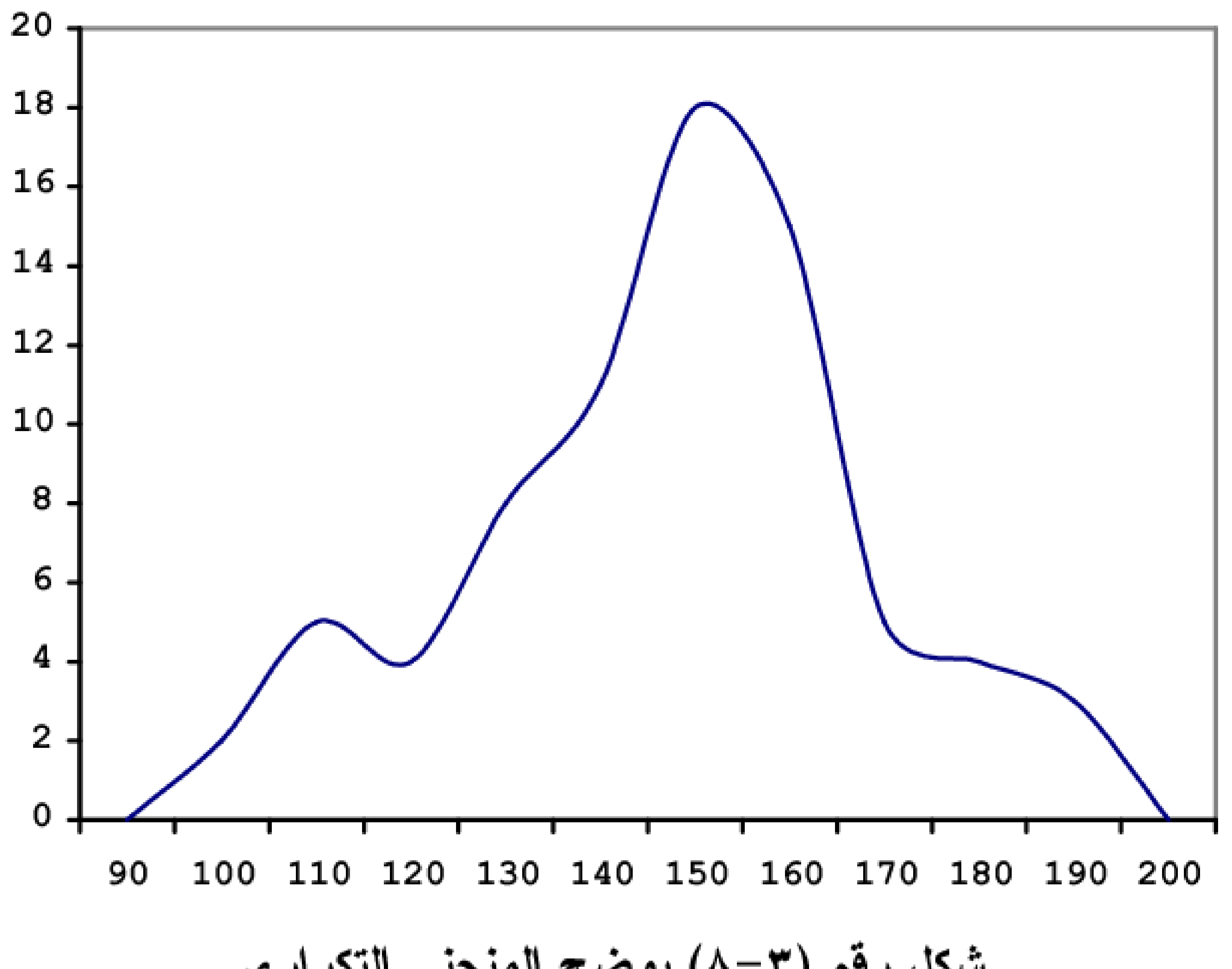
الأعلـــى لكل فئة وبتوصيل تلك القيم بعضها ببعض باستخدام خطوط مستقيمة فإننا نحصل على المضلع التكرارى التجمعى الموضح بالشكل (٣- ٧).



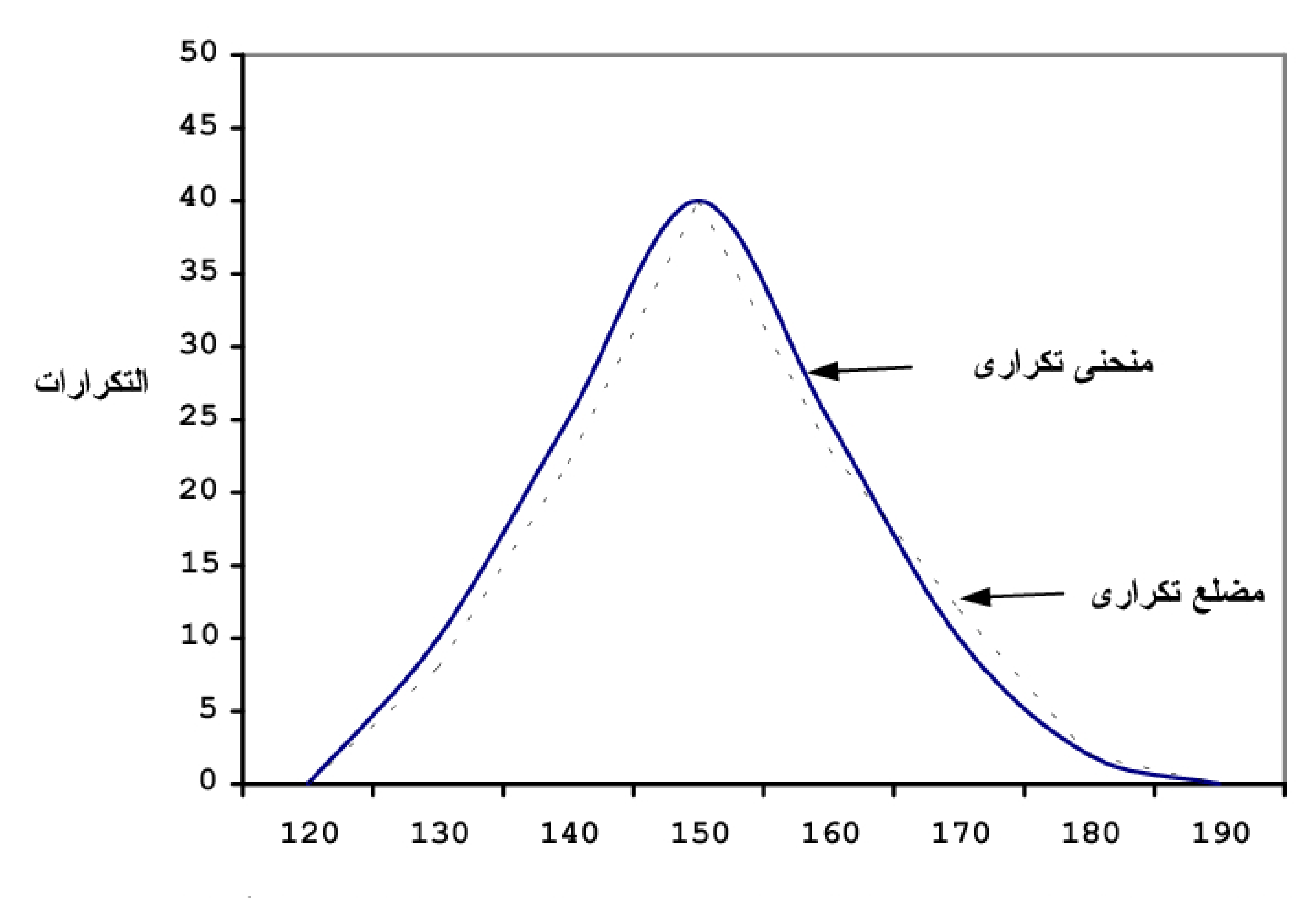
شكل رقم (٣- ٧) المضلع التكرارى التجمعى للأوزان

٢- المنحنى التكرارى Frequency Curve :

فى بعض الأحيان يرغب الباحثون فى التخلص من التعرجات أو الانكسارات الــتى يتصف بها المضلع التكرارى والتوصل إلى شكل أكثر سلاسة Smoothing وتحويل المضلع إلى منحنى تكرارى. ومن ثم يمكن القول أن المنحنى التكرارى لا يختلف عن المضلع التكرارى سواءً من حيث الشكل أو طريقة الرسم. فإذا كنا نقوم بتوصــيل الــنقط المتتالية بخطوط مستقيمة تصل بينها فى المضلع التكرارى، ففى حالــة المنحنى، نقوم باستخدام اليد بتوصيل النقط القريبة من بعضها فى القيم دون الاهتمام بالنقط القريبة منها والشاذة فى قيمتها أحياناً، سواءً أكانت تلك النقط الشاذة تعلــو أو تقــل عن منحنى التوصيل بين النقط القريبة (ويوضح الشكل رقم ٣-٨) المنحــنى التكرارى. ومن ذلك لا يمكن القول إن المساحة تحت المنحنى التكرارى تساوى المساحة تحت المضلع التكرارى بل ستكون المساحة الأولى أقل من الثانية كما يتضح ذلك من الشكل رقم (٣-٨).



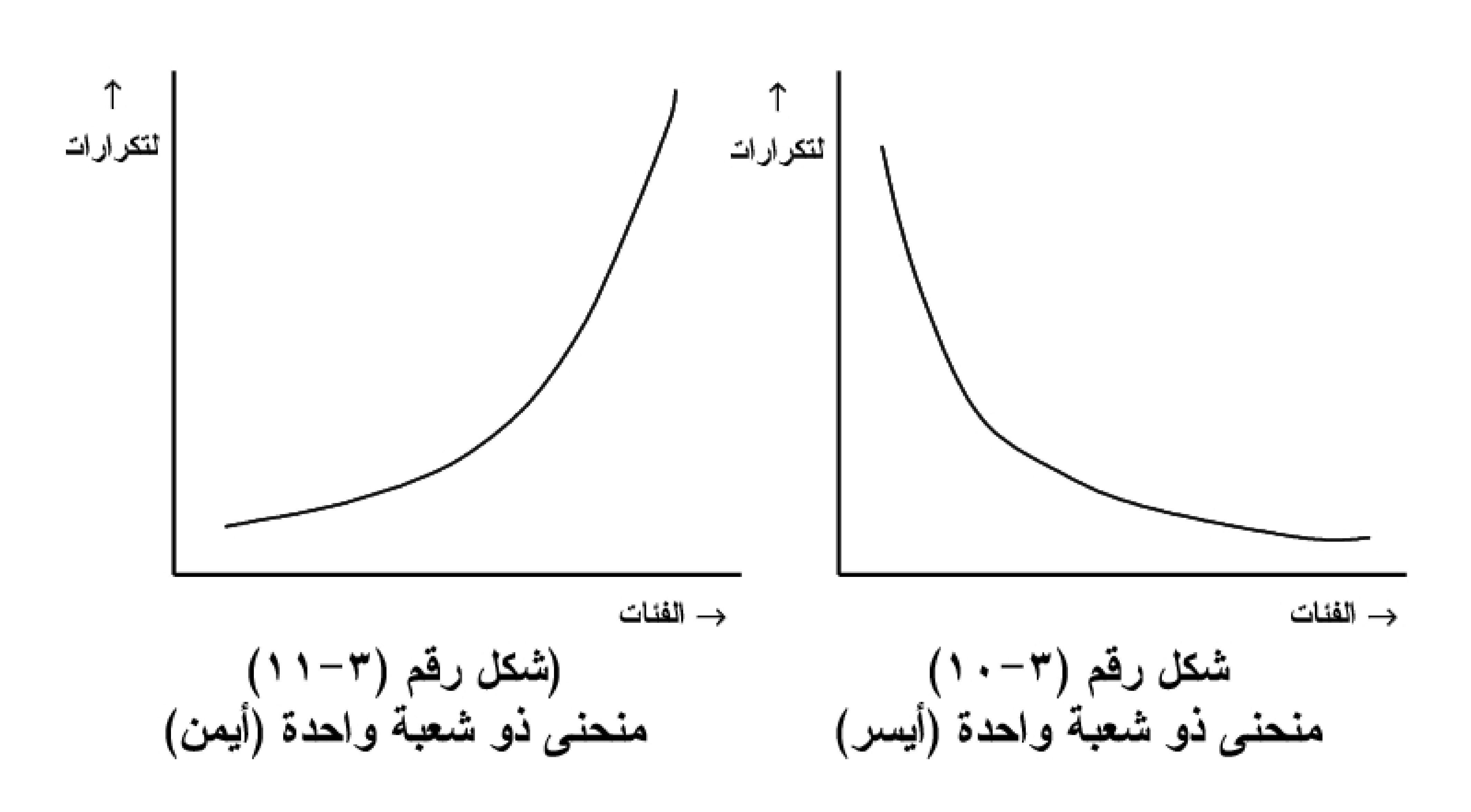
شكل رقم (٣-٨) يوضح المنحنى التكرارى



شكل رقم (٣-٩) رسم المنحنى التكرارى من المضلع التكرارى للأوزان

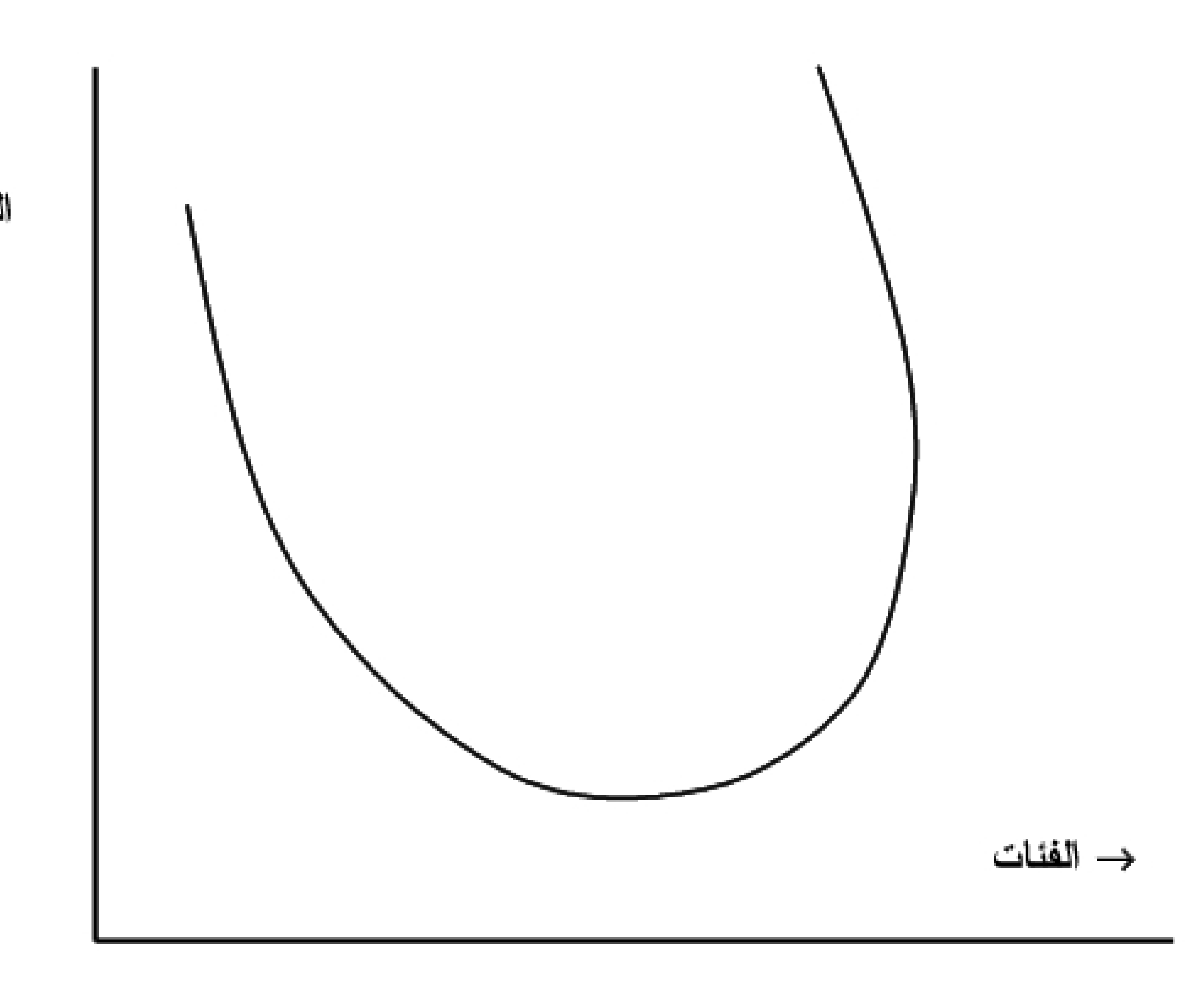
ومن خصائص المنحنيات التكرارية – كما تناولناها سابقًا بالشرح – خاصيتى الالتواء والتفلطح وأيضًا المنحنى الاعتدالى. وفى بعض الحالات التى تقل فيها عدد التكرارت، بينما تزيد أعداد أخرى بشكل لا يقارن فإن المنحنى التكرارى يكون فى هـذه الحالة ذا شعبة واحدة. مثال ذلك توزيع السكان على أساس الثروة، حيث نجد الغالبـية العظمـى مـن الفقـراء، بينما قلة قليلة جدًا من الأغنياء. وبالمثل توزيع أراضـى الإصـلاح الـزراعى، فالحيازة قليلة وعدد الحائزين كبير، فى حين تقل أعداد الملاك الزراعيين أصحاب الأراضى الزراعية كبيرة المساحة. وأمثلة أخرى مشابهة نجد أن المنحنى التكرارى ذو شعبة واحدة كما هو موضح فى الشكلين رقم (٣-١٠)، (٣-١١).

وفـى حالة تجمع التكرارات الكبيرة نسبيًا عند طرفى المنحنى كما هو الحال فـى تعداد الوفيات على مستوى الجمهورية، حيث تكثر نسبة الوفيات فى مرحلتى الشيخوخة والطفولة المبكرة خاصة السنتين الأولى والثانية من عمر الأطفال، بينما يقـل عـدد الوفيات نسبيًا وبدرجة كبيرة بين الفئة الشبابية، وفى هذه الحالة نجد أن المنحـنى من النوع الناقوسى المقلوب، وليس شرطًا أن يكون اعتداليًا ويسمى عليه المنحـنى التكرارى ذو الشعبتين كما يوضحه الشكل رقم (٣-١٢) ويمكن ملاحظة أن تكرارت الفئات الوسطى تقل كثيرًا عن تكرارات فئات الطرفين.



شكل رقم (٣-١٠)
منحنى ذو شعبة واحدة (أيسر)

(شكل رقم (٣-١١)
منحنى ذو شعبة واحدة (أيمن)



شكل رقم (٣-١٢) منحنى ذو الشعبتين

المنحنيات المتجمعة Cumulative :

تعتبر المنحنيات المتجمعة التمثيل البيانى الرابع للتوزيعات التكرارية ويقتصر استخدامها فى حالات تجميع التكرارات فى فئات متتالية. والمنحنيات المتجمعة نوعان إما هابطة أو صاعدة. وتستخدم المنحنيات المتجمعة بنوعيها لمعرفة عدد أو نسبة المفردات التى تزيد أو تقل عن قيمة أو نسبة معينة، أو إذا أردنا معرفة الوضع النسبى لقيمة معينة من قيم المتغير.

خطوات رسم المنحنى المتجمع بنوعيه الصاعد والهابط:

١- استخدام المحاور الإحداثية (س، ص) فى الشكل البيانى بأن يكون المحور السينى مقسمًا لفئات متساوية الأبعاد كذلك تخصيص المحور الصادى للتكرارات أو التكرارات النسبية.

٢- يتم توصيل النقط التى تمثل التكرارات المتجمعة عند الحدود العليا للفئات فنحصل على خط ممهد يمثل المنحنى المتجمع الصاعد ولتحقيق ذلك لابد أن نقوم أولاً بتحويل جدول التكرارات المعطى إلى جدول تكرارى صاعد. ولو كانت التكرارات المتجمعة عند أى فئة تمثل التكرارات التى تزيد عن الحد الأدنى لتلك الفئة، فإذا قمنا بتوصيل النقط التى تمثل التكرارات المتجمعة عند الحدود الدنيا للفئات فسنحصل على المنحنى المتجمع الهابط. ولتحقيق ذلك أيضًا يجب على الدارس تحويل جدول التوزيع التكرارى المعطى له إلى

جـدول مـتجمع هابط بنفس الخطوات السابق شرحها فى عمل جدول التجمع الهابط.

٣- لا يشـترط التقـيد بـنفس مقياس الرسم للمحورين السينى والصادى، كما لا يشـترط أن يبدأ تقسيم المحور السينى من نقطة الأصل بالقيمة صفر، لتفادى كـبر حجـم التمثيل البيانى. أما فى حالة رسم المنحنيين الصاعد والهابط معًا علـى المحـاور الإحداثـية فيفضل أن تستخدم لهما نفس مقياس الرسم حيث يلتقيان معًا فى نقطة يساوى إحداثيها الصادى نصف مجموع التكرارات.

٤- فـى حالـة عـدم تسـاوى طـول الفئات، لا يقوم الدارس بإجراء أى تعديل للـتكرارات بـل يلاحظ فقط صحة رصد القيم التكرارية لحدى الفئة (الأدنى والأعلـى). والسـبب فـى عدم الحاجة لتعديل التكرارات أن التمثيل بمنحنى متجمع لا يعدو كونه عملية تجميع فقط.

مثال:

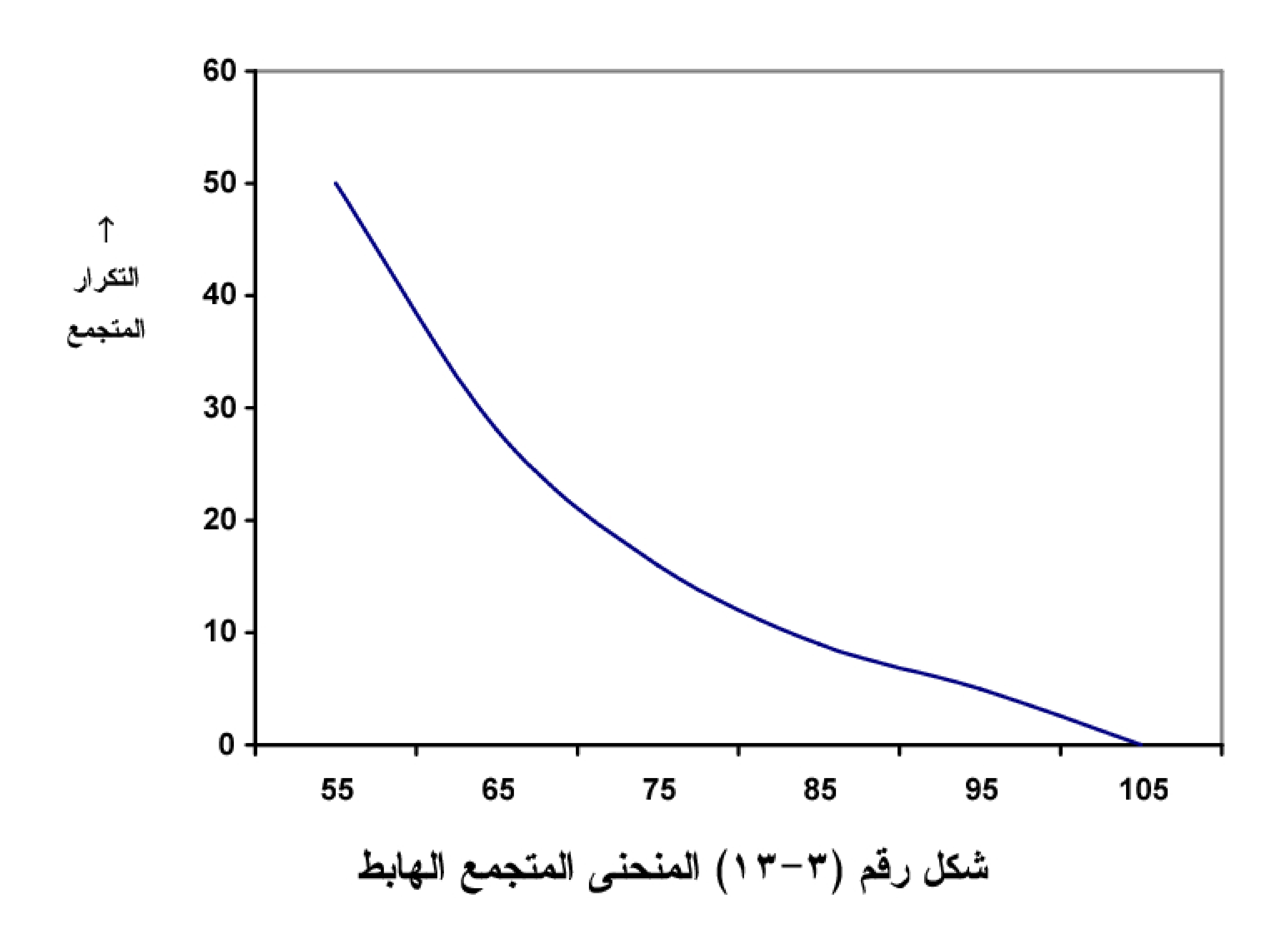
ارسم المنحنيين المتجمعين الصاعد والهابط من الجدول التكرارى التالى.

ف	ك
٥٥-	٢٢
٦٥-	١٢
٧٥-	٧
٨٥-	٤
٩٥-١٠٥	٥
المجموع	٥٠

الحـل:

جدول متجمع هابط

فئات	ك	الحدود الدنيا للفئات	التكرار المتجمع الهابط
٥٥-	٢٢	٥٥ فأكثر	٥٠
٦٥-	١٢	٦٥ فأكثر	٢٨
٧٥-	٧	٧٥ فأكثر	١٦
٨٥-	٤	٨٥ فأكثر	٩
٩٥-١٠٥	٥	٩٥ فأكثر	٥
مجـ	٥٠		

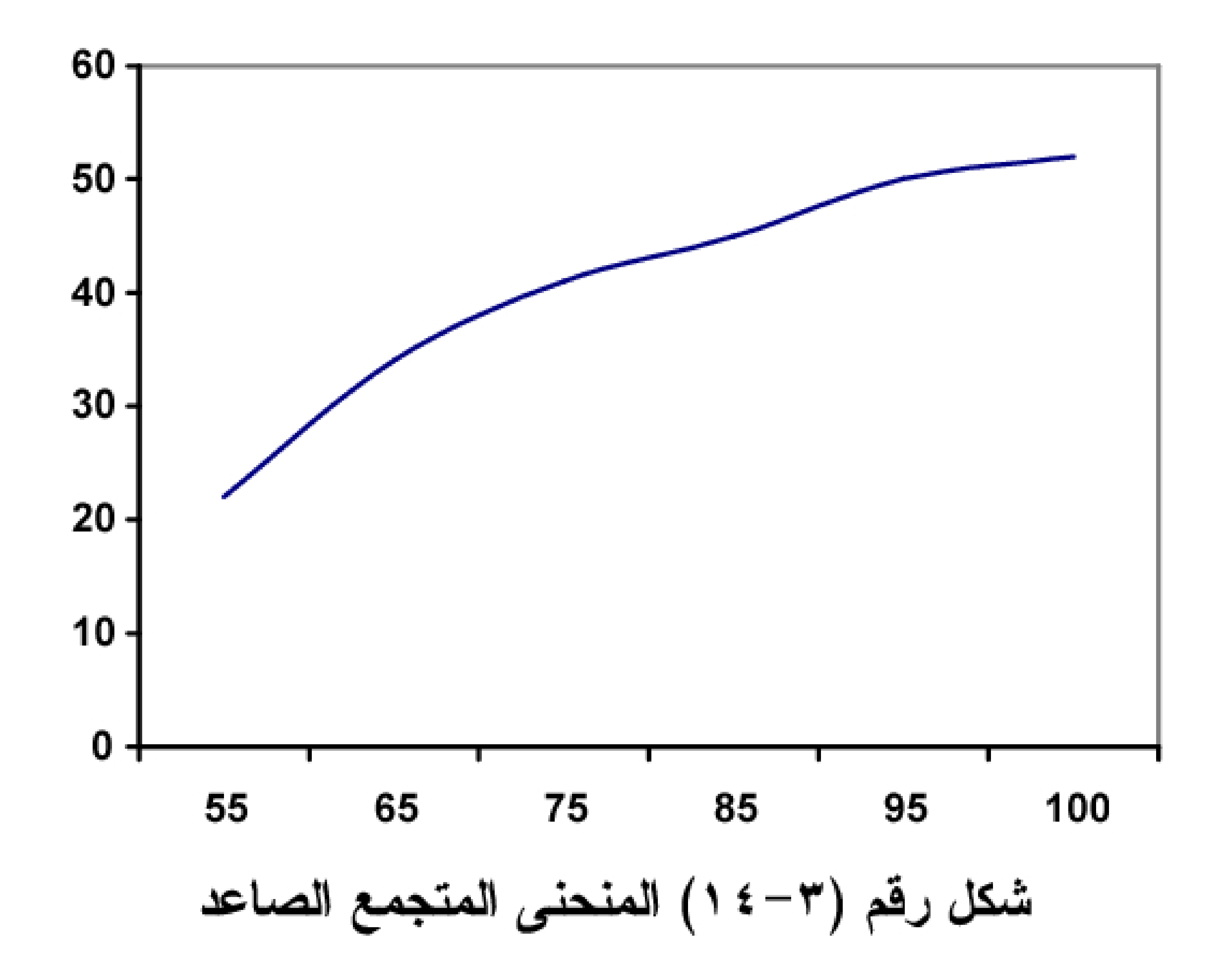


شكل رقم (٣-١٣) المنحنى المتجمع الهابط

جدول متجمع صاعد

فئات	ك	الحدود العليا للفئات	التكرار المتجمع الصاعد
٥٥-	٢٢	أقل من ٦٥	٢٢
٦٥-	١٢	أقل من ٧٥	٣٤
٧٥-	٧	أقل من ٨٥	٤١
٨٥-	٤	أقل من ٩٥	٤٥
٩٥-١٠٥	٥	أقل من ١٠٥	٥٠
مجـ	٥٠		

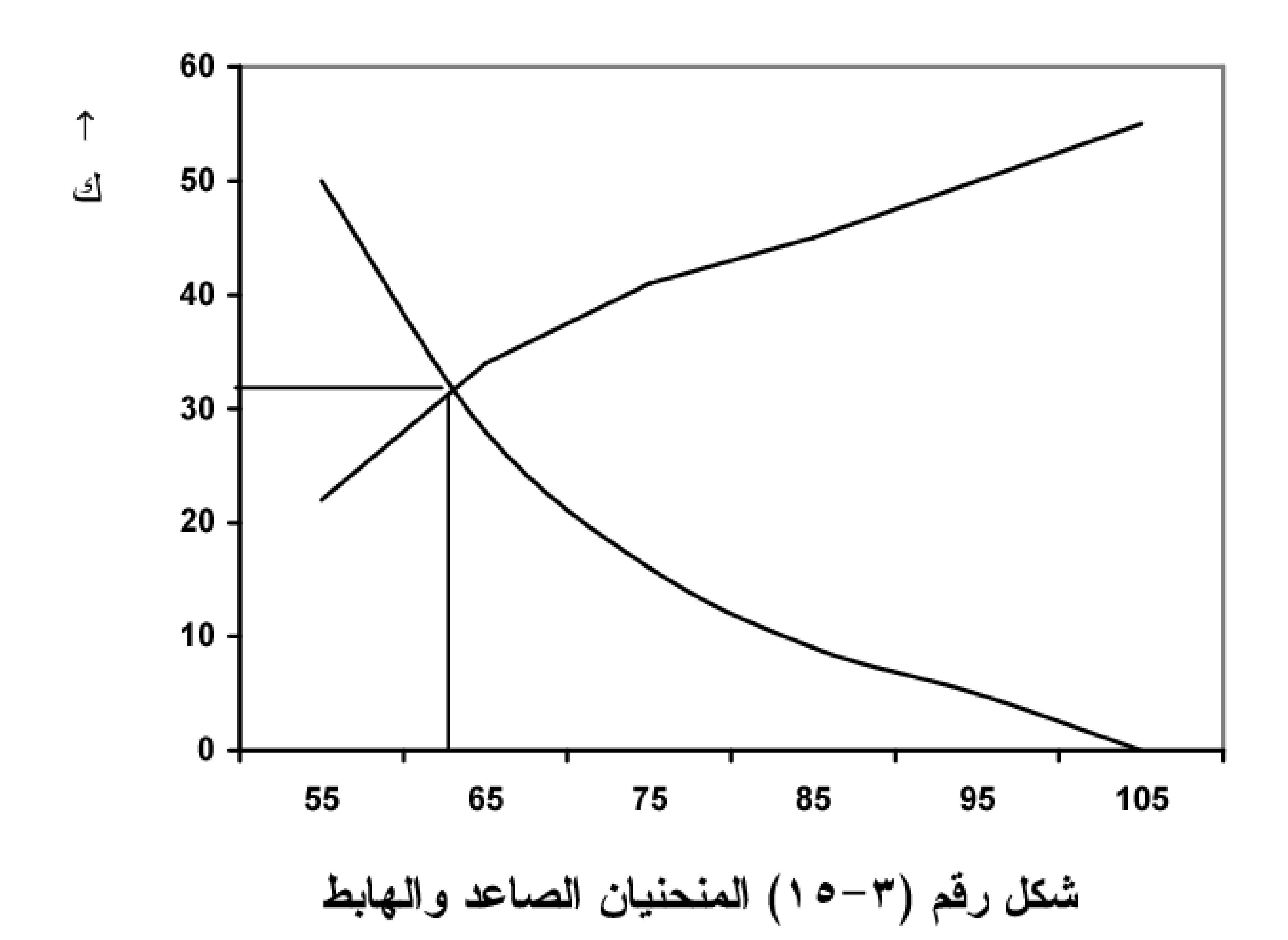
لاحـظ أن تكـرار الفـئة الأخيرة للتكرار المتجمع الصاعد يساوى فى القيمة مجموع التكرارات الأصلية.



شكل رقم (٣-١٤) المنحنى المتجمع الصاعد

ويمكن رسم المنحنيين فى شكل واحد كما يتضح من الشكل رقم (٣-١٥)

لاحــظ أن نقطــة تقــاطع المنحنيين الصاعد والهابط عند تكرار (٢٥) على المحور الصادى وهى قيمة تمثل نصف مجموع التكرارات.



شكل رقم (٣-١٥) المنحنيان الصاعد والهابط

المفاهيم الأساسية Key Concepts

جداول التوزيع التكرارى النسبى Percentage Frequency Table:

يقصـــد بالتكرار النسبى لفئة ما هو تكرارها بالنسبة إلى التكرار الكلى لجميع الفئات.

جداول التكرار التجمعى Cumulative Freuqncy Tables:

يســـتخدم هـــذا الـــنوع مـــن الجداول التكرارية إذا أراد باحث أن يعرف عدد المفردات التى تقل أو تزيد عن قيمة معينة.

الجداول المزدوجة:

تســـتخدم فـــى تلخـــيص ازدواج القيم لمتغيرين حيث يتم تبويب البيانات وفقًا لفئتيـــن فـــى ترتيب صفوف وأعمدة بحيث تشمل الصفوف تكرارات الصفة الأولى بينما تشمل الأعمدة تكرارات المتغير الثانى.

تمارين الفصلين الثانى والثالث

١- باستخدام الدائرة وضح النسبة المئوية للإنفاق على مجالات الرعاية الصحية الموضحة، والتى تستنفد ميزانية وزارة الصحة وذلك من الجدول الإحصائى التالى :

مجال الإنفاق	النسبة المئوية
رعاية علاجية بالمستشفيات	٣١,٤٠
خدمات طبية	٢٧,٦٩
خدمات أسنان	١٠,٨٤
خدمات مهنية طبية أخرى	٢,٨٤
أدوية ومستلزماتها	١٢,٩٥
نظارات وعدسات طبية	٢,٢٣
رعاية تمريض بالمنزل	٦,٠٩
مجالات أخرى متعلقة بالرعاية الصحية	٥,٩٦

٢- بدءًا من إعلان سياسة الانفتاح عام ١٩٧٦، أخذت فى التزايد موجات هجرة المصريين خاصة للدول النفطية العربية. ولقد أسفرت إحدى الدراسات التى أجريت على إحدى مراكز الوجه البحرى لمعرفة النسب المئوية للمهاجرين خلال الفترة من عام ١٩٧٦ حتى عام ١٩٨١ عن النتائج التالية :

السنة	١٩٧٦	١٩٧٧	١٩٧٨	١٩٧٩	١٩٨٠	١٩٨١
النسبة المئوية للمهاجرين بالنسبة للسكان بالمركز	٣,٤	٥,٥	٦,١	٣,٤	٨,٨	١٢,٥

(أ) هل يعد استخدام الدائرة أفضل فى تمثيل البيانات لنسب المهاجرين؟

(ب) استخدم المستطيلات فى التمثيل البيانى للبيانات السابقة.

٣- فيما يلى أرقام افتراضية عن عدد الأسر المستفيدة من الوحدات الصحية ببعض القرى والمطلوب عرضها بيانيًا.

البيان	١٩٧٠	١٩٨٠
أ	٢٥٧١	٥٢٦٧
ب	٤٠٧٦	٩٨٨١
ج	٥٩٥٨	٩٥١٥
د	١٥٣٤	٧٧٢١
هـ	٢٩٨٢	٣٧٨٩

٤- فـيما يلى عدد السكان فى مدينة ما موزعة ذكورًا وإناثًا والمطلوب عرضها بيانيًا بالأساليب التى تراها مناسبة :

السنة	ذكور	إناث	المجموع
١٩٨٠	١٧٩٣٨	٦٣٦٢	٢٤٣٠٠
١٩٨١	١٥٦١١	٥٧٠٠	٢١٣١١
١٩٨٢	١٤٨٤٢	٥٨٨٧	٢٠٧٢٩
١٩٨٣	١٥٢٤٥	٦٨٠١	٢٢٠٤٦

٥- فيما يلى درجات لعينة من الطلاب فى أحد الامتحانات :

٩٩	٨٨	٨٢	٧٧	٧٤	٦٩
٩٦	٨٧	٨٠	٧٧	٧٣	٦٩
٩٣	٨٦	٨٠	٧٦	٧٣	٦٧
٩١	٨٤	٧	٧٤	٧١	٦٦
٩٠	٨٣	٧٨	٧٤	٧٠	٦٤
٩٠	٨٢	٧٧	٧٤	٧٠	٦١

(أ) ابدأ بعمل الفئة ٦٠ – ٦٢ واستكمل الفئات لتشمل جميع الدرجات ثم ارسم المدرج التكرارى لتلك البيانات.

(ب) استخدم نفس الفئات لرسم المضلع التكرارى.

(جـ) صِف توزيع تلك القيم.

(د) صـنف الدرجـات السـابقة فـى جدول تكرارى مدى كل فئة فيه خمس درجات.

٦- تمـــثل البيانات التالية الحالة الزواجية لعينة، تم أخذها من أحد الأحياء السكنية بمديـــنة القاهرة، والمطلوب تنظيم هذه البيانات، وعمل جدول توزيع تكرارى للحالة الزواجية لأفراد العينة.

متزوج	مطلق	متزوج	متزوج	أرمل	متزوج
أرمل	متزوج	متزوج	متزوج	متزوج	متزوج
متزوج	متزوج	متزوج	مطلق	أعزب	متزوج
متزوج	أعزب	متزوج	أعزب	أعزب	متزوج
متزوج	أعزب	أعزب	أعزب	أعزب	أعزب
أرمل	أعزب	أعزب	أعزب	أعزب	مطلق
أعزب	مطلق	عزب	متزوج	أعزب	أعزب
مطلق	أرمل	أرمل	متزوج	أعزب	أعزب
أرمل	متزوج	متزوج	متزوج	متزوج	متزوج
متزوج	متزوج	متزوج	متزوج	متزوج	متزوج

٧- أكمل ما يأتى :

(أ) يفضـــل اســـتخدام النســـب المئوية أو النسب بدلاً من التكرارات فى عمل ....

(ب) تعتـــبر الأعمدة البيانية والدائرة بقطاعاتها أفضل أساليب الرسومات البيانية للبيانات ....

(ج) يعتبر المضلع شكلاً بيانيًا ملائمًا لمستوى .... من المتغيرات، ويكـــون وضـــع القـــيم علـــى المحور السينى (الأفقى) باستخدام .... للتوزيع التكرارى.

(د) يستخدم المضلع التكرارى نقطة عند وسط كل فئة بدلاً من .... لتمثيل التكرارات.

٨- فـــيما يلى عدد من الاختيارات قد تكون إجابة أو أكثر هى الإجابة الصحيحة الـــتى تكمل العبارة فى السؤال ذاته. والمطلوب وضع علامة دائرة (O) على رقم الاختيار الصحيح.

أ - إذا قمـــت برســـم خطوط مستقيمة تربط بين منتصف الأعمدة للمضلع، سوف تحصل على :

١- منحنى اعتيادى.

٢- توزيع تكرارى.

٣- مضلع تكرارى.

٤- ستحدث مشكلة كبرى.

ب- عند عمل توزيع تكرارى فإن أول خطوة يجب عملها هى:

١- إيجاد مدى القيم.

٢- أخذ الحدود الحقيقة.

٣- تحديد مدى الفئة.

٤- حساب الفئة المئوية للتوزيع.

٥- لا إجابة من الإجابات الأربع السابقة.

ج- إن النقاط التى أقوم بالتوصيل بينها بخطوط فى المضلع التكرارى تتطابق مع:

١- الحد الأدنى للفئة.

٢- الحد الأعلى للفئة.

٣- النقطة المتوسطة للفئة.

٤- النقاط النهائية للفئة.

٥- ليس للفئة أى حدود.

٤

الفصل الرابع
مقاييس النزعة المركزية

مقدمة
أولاً: المتوسط الحسابى.
ثانيًا: الوسيط.
ثالثًا: المنوال.

الفصل الرابع
مقاييس النزعة المركزية

مقدمة :

يتـــناول هـــذا الفصـــل بالشـــرح عددًا من مقاييس النزعة المركزية الشائعة الاســـتخدام فـــى البحوث الاجتماعية، وتتمثل هذه المقاييس فى المتوسط الحسابى، الوســـيط والمـــنوال. وسوف نعرض لكيفية حساب كل منها سواء بالنسبة للبيانات الخام (غير المُبَوَّبة) أو البيانات المُبَوَّبة.

ويهـــدف الفصل إلى أن يعرف الطالب حساب كل مقياس من مقاييس النزعة المركـــزية: المتوســـط الحســـابى والوســـيط والمنوال، للبيانات الخام أو للجداول الـــتكرارية. وأن يقـــارن بيـــن المقايـــيس الثلاثة ويعرف مزايا وعيوب كل منها. ويستطيع أن يوجد قيمة كل من المنوال والوسيط بالرسم.

أولاً: المتوسط الحسابى :

١ – حساب المتوسط الحسابى من البيانات غير المبوبة:

يعتـــبر المتوســـط والوســـيط أهـــم مقاييس النزعة المركزية التى تستخدم فى الـــبحوث الاجتماعية، وإن كان المتوسط الحسابى أكثرهما شيوعًا. ولعل السبب فى ذلـــك يـــرجع إلـــى أن المتوســـط الحسابى هو أصدق هذه المقاييس الثلاثة تمثيلاً للمجموعة موضوع الدراسة، وبالتالى تصبح تلك المجموعة مؤشرًا جيدًا للمجموعة الأصلية نظرًا لأن المتوسط دائمًا ما يكون من نفس وحدات المتغير. ومن هنا يمكن تعريف المتوسط بأنه حاصل قسمة مجموع القيم على المجموع الكلى لأفراد العينة. ومـــن ثـــم لـــو افترضـــنا أن القيم هى س١، س٢، س٣ ... حتى س ن وأن عدد مفـــردات العيـــنة هى (ن) فإن قيمة المتوسط الحسابى ويرمز له ($\overline{\text{س}}$) يتم حسابها من المعادلة التالية :

$$\overline{\text{س}} = \frac{\text{س}_{١} + \text{س}_{٢} + \text{س}_{٣} + \ldots\ldots + \text{س}_{\text{ن}}}{\text{ن}}$$

$$\boxed{\overline{\text{س}} = \frac{\text{مجـ س}}{\text{ن}}}$$

حيث $س_1$ تمثل قيمة المفردة الأولى، و$س_2$ تمثل قيمة المفردة الثانية وهكذا حتى $س_ن$ قيمة المفردة الأخيرة.

مثال: إذا كانت ١٢، ١٨، ١٠، ٧ هى الدرجات التى حصلت عليها أربع طالبات فى اختبار نصف العام لمادة الإحصاء الاجتماعى. احسب المتوسط الحسابى لتلك القيم.

$$\therefore \bar{س} = \frac{١٢ + ١٨ + ١٠ + ٧}{٤} = ١١,٧٥$$

ويتصف المتوسط الحسابى بخاصية جبرية أساسية وهى أن مجموع انحرافات القيم عن المتوسط الحسابى لها لابد أن يساوى صفرًا. أى أن:

$$مجـ (س - \bar{س}) = صفرًا$$

ولا غرابة فى اتصاف المتوسط الحسابى بتلك الخاصية، والتى يمكن استنباطها منطقيًا من التعريف السابق، كما يمكن إثباتها من المثال التالى:

مثال: نفترض أننا نريد حساب المتوسط من الأرقام التالية:

٧٢ ، ٨١، ٨٦، ٦٩ ، ٥٧.

الحل:

$$\text{من العلاقة السابقة } \bar{س} = \frac{٧٢ + ٨١ + ٨٦ + ٦٩ + ٥٧}{٥} = ٧٣$$

فلو قمنا بطرح قيمة $\bar{س}$ وهى (٧٣) من كل قيمة من القيم الست سيكون ناتج جمع الفروق يساوى صفرًا كما يتضح ذلك من الجدول التالى :

جدول (٤-١)

س	س - ٧٣
٧٢	-١
٨١	٨
٨٦	١٣
٦٩	-٤
٥٧	-١٦
مجـ	صفر

٢- حساب المتوسط الحسابى من البيانات المبوبة Grouped Data :

أ - حساب المتوسط الحسابى من خلال التوزيع التكرارى البسيط.

مثال: فيما يلى توزيع الأجر اليومى (بالجنيه المصرى) (س) لعدد من عمال الخدمات.

الأجر	٥	٧	٨	١١	١٢	١٤	١٥
التكرار	٣	٢	٩	١٥	٦	٣	٢

الحـــل :

جدول (٤-٢)

س	ك	س ك
٥	٣	١٥
٧	٢	١٤
٨	٩	٧٢
١١	١٥	١٦٥
١٢	٦	٧٢
١٤	٣	٤٢
١٥	٢	٣٠
	٤٠	٤١٠

$$\overline{س} = \frac{مجـ\ س\ ك}{مجـ\ ك}$$

$$= \frac{٤١٠}{٤٠} = ١٠,٢٥ \text{ جنيهًا}$$

ب- حساب المتوسط الحسابى من التوزيع التكرارى ذى الفئات:

فيما يلى توزيع تكرارى لدرجات عينة من الطلاب فى امتحان مادة الإحصاء والمطلوب حساب المتوسط الحسابى لهذه المجموعة.

جدول رقم (٤-٣)

فئات	ك
٢٠-	١٢
٣٠-	٨
٤٠-	٢
٥٠-	٩
٦٠-	٥٠
٧٠-	١٣
٨٠-	٧
٩٠-١٠٠	٨
مجـ	١٠٩

خطوات الحل :

١- أوجد مركز كل فئة (س).
٢- اضرب مركز الفئة فى التكرار المناظر للفئة ذاتها (س ك).
٣- طبق المعادلة الآتية :

$$\overline{س} = \frac{مجـ\ س \times ك}{مجـ\ ك}$$

جدول رقم (٤-٤)

ف	ك	س	س ك
٢٠-	١٢	٢٥	٣٠٠
٣٠-	٨	٣٥	٢٨٠
٤٠-	٢	٤٥	٩٠
٥٠-	٩	٥٥	٤٩٥
٦٠-	٥٠	٦٥	٣٢٥٠
٧٠-	١٣	٧٥	٩٧٥
٨٠-	٧	٨٥	٥٩٥
٩٠-١٠٠	٨	٩٥	٧٦٠
	١٠٩		٦٧٤٥

نســـتخلص مما سبق إلى أنه فى حالة (المتغير الواحد) فى البيانات التكرارية المجدولـــة مثل الطول، الأجر أو الوزن فإن قيمة المتوسط الحسابى ($\overline{س}$) تحسب من العلاقة التالية :

$$\overline{س} = \frac{مجـ\ س \times ك}{مجـ\ ك}$$

$$= \frac{٦٧٤٥}{١٠٩} = ٦١,٩$$

أيضًا فى حالة البيانات غير المبوبة مثل سلسلة من الأرقام كما ذكرنا فى مثال ســابق أو تســجيل عدد من الأرقام فى إحدى التمارين الرياضية يحسب المتوسط الحسابى ($\overline{س}$) من العلاقة التالية :

$$\overline{س} = \frac{مجـ\ س}{ن}$$

٣- المتوسط المرجح Weighted Mean :

إذا كـــان جـــدول الـــتوزيع الـــتكرارى مفـــتوح النهاية كما فى الجدول رقم (٤-٥) حجـــم الوحـــدة المعيشية ستة أفراد فأكثر (٦+) بمعنى أن حجم الوحدة قد يكـــون ٦، ٧، ٨، ٩، .... إلـــخ. ولحساب المتوسط المرجح فى هذه الحالة نختار رقمًـــا أكـــبر من (٦) وليكن رقم (٩) من ثم سوف يتغير مجموع (س ك) وبالتالى قيمة المتوسط الحسابى المرجح.

جدول رقم (٤-٥)

التوزيع التكرارى لعينة من الوحدات المعيشية وفقًا للحجم فى عامى ١٩٩٨، ٢٠٠٨

س	١٩٩٨		٢٠٠٨	
	ك	س ك	ك	س ك
١	١٢	١٢	٢٥	٢٥
٢	٣٠	٦٠	٣٤	٦٨
٣	٢٣	٦٩	١٧	٥١
٤	١٩	٧٦	١٦	٦٤
٥	٩	٤٥	٦	٣٠
٦فأكثر افتراضيًا=٩	٧	٦٣	٢	١٨
مجـــ	١٠٠	٣٢٥	١٠٠	٢٥٦

∴ متوسط حجم الوحدة المعيشية فى عام ١٩٩٨ = $\frac{٣٢٥}{١٠٠}$

= ٣,٢٥ فردًا لكل أسرة

المتوسط المرجح لحجم الأسر فى عام ٢٠٠٨ = $\frac{٢٥٦}{١٠٠}$

= ٢,٥٦ فردًا لكل أسرة

٤– متوسط الجماعات المشتركة The Mean of Combined Groups:

يستخدم متوسط الجماعة فى البيانات المبوبة التى تشتمل على جماعات مشتركة فى ظاهرة معينة رغم أنها تختلف فيما بينها من حيث عدد المفردات والنوعية. ويوضح المثال التالى كيف يمكن حساب متوسط المجموعة للجماعات المشتركة.

مثال :

فى امتحان مادة الأنثروبولوجيا الحضرية، كان عدد الطلاب الذين تقدموا للامتحان ١٥٥ وكان عدد الطالبات ١٠٨ طالبة، وعدد الطلاب ٤٧ طالبًا $\bar{س}$ أ = ٤٥,٢٦ فى حين بلغ المتوسط الحسابى لدرجات الطلاب، $\bar{س}$ ذ = ٥٤,٨٩. والمطلوب حساب المتوسط العام للطالبات والطلاب داخل الجماعة (م) (أو متوسط التوزيع الكلى للطلاب والطالبات).

الحـل :

$$م/ = \frac{(مجـ ك_ذ \ \bar{س}_ذ) + (مجـ ك_أ \ \bar{س}_أ)}{مجـ ك_ذ + مجـ ك_أ}$$

حيث إن :

مجـ ك $_أ$ = مجموع تكرارات الإناث.

مجـ ك $_ذ$ = مجموع تكرارات الذكور.

$\bar{س}$ $_أ$ = المتوسط الحسابى للإناث

$\bar{س}$ $_ذ$ = المتوسط الحسابى للذكور.

$$\therefore \text{المتوسط} = \frac{(٤٧)\ (٥٤,٨٩) + (١٠٨)\ (٤٥,٢٦)}{٤٧ + ١٠٨}$$

= ٤٨,١٨ درجة

جدول رقم (٤-٦)
التوزيع التكرارى لدرجات الامتحان النهائى
لمادة الأنثروبولوجيا الحضرية

الدرجة (س)	ك
٧٩	٢
٦٨	١
٦٥	٥
٦٣	١٢
٦٠	١٧
٥٨	١١
٥٣	٢٠
٥٢	١٣
٥١	٢٣
٤٧	١٨
٤٥	١٤
٤٠	١١
٣٨	٥
٣٥	٣
مجـ	١٥٥

ثانيًا : الوسيط :

يعتبر الوسيط مقياسًا ترتيبيًا Ordinal، على عكس المتوسط الحسابى والذى لا يستخدم إلا كمقياس كمى Interval Measurement، ويُعرف الوسيط بأنه النقطة أو القيمة التى تقسم القيم التجريبية أو المشاهدات التى تسجل حول ظاهرة ما إلى مجموعتين، شريطة أن يتساوى عدد القيم الأكبر عن الوسيط مع باقى القيم الأصغر منه والتى تليه فى الترتيب (بمعنى أن ٥٠% من القيم أكبر من قيمة

الوسيط، و ٥٠% من القيم أقل من قيمة الوسيط)، حيث يتم ترتيب تلك القيم جميعها أما ترتيبًا تصاعديًا أو تنازليًا.

ويتصــف الوسيط بخاصية مهمة وهى عدم تأثره بالقيم المتطرفة على جانبى منحــنى التوزيعات للقيم المبوبة. ومن ثم يفضل فى تلك الحالات استخدام المتوسط الحســابى. ويمكن استخدام الوسيط فى حالة التوزيعات التكرارية التى تتصف بحدة الالــتواء وأيضًا فى حالة التوزيعات التكرارية المفتوحة. وهذه خاصية مهمة يتميز بها الوسيط عن كل من المتوسط الحسابى والمنوال معًا.

حساب الوسيط فى حالة البيانات غير المبوبة (البيانات الخام):

١- إذا كان عدد القيم فرديًا :

يــتم أولاً ترتيب القيم تصاعديًا أو تنازليًا، فإذا كان عدد تلك القيم فرديًا تكون قــيمة الوسيط هو حاصل قسمة هذا العدد إجمالاً مضافًا إليه الواحد الصحيح على (٢). فلو فرضنا أن عدد القيم (ن).

$$\text{الوسيط} = \frac{\text{ن} + ١}{٢}$$

والناتج من المعادلة يمثل موقع قيمة الوسيط داخل ترتيب كل القيم.

مــثال: حصــل ١٣ طالــبًا على الدرجات التالية فى اختبار أعمال السنة فى مادة الاجتماع الصناعى. احسب الوسيط لتلك الدرجات :

٤ ، ٩ ، ١٥ ، ٨ ، ١٤ ، ٦ ، ٥ ، ٧ ، ٩ ، ١٢ ، ١٠ ، ١٣ ، ١٦.

الحــل :

يــتم أولاً ترتيــب الدرجات المعطاة وفقًا لقيمها ترتيبًا تنازليًا أو تصاعديًا كما قلنا. وفى هذا المثال اخترنا الترتيب التصاعدى.

٤ ، ٥ ، ٦ ، ٧ ، ٨ ، ٩ ، [٩] ، ١٠ ، ١٢ ، ١٣ ، ١٤ ، ١٥ ، ١٦.

$$\text{الوسيــط} = \frac{١٣ + ١}{٢}$$

$$= \frac{١٤}{٢} = ٧ \text{ أى القيمة السابعة}$$

أى أن الوسيط هو القيمة السابعة فى الترتيب التصاعدى وهى ٩. حيث توجد (٦) قـيم سابقة هى ٤ ، ٥ ، ٦، ٧، ٨، ٩ وأيضًا ست قيم لاحقة لقيمة الوسيط هى ١٠، ١٢، ١٣، ١٤، ١٥، ١٦.

ب- إذا كان عدد القيم زوجيًا:

إذا كان عدد القيم زوجيًا، توجد قيمتين وسيطتين الأولى عند $\frac{(ن)}{٢}$ والثانـية عند $(\frac{ن}{٢}) + ١$

مـثال: فـيما يلى درجات عشرة طلاب حصلوا عليها فى اختبار آخر العام لمادة النصوص الأجنبية فى قسم الاجتماع. والمطلوب حساب الوسيط.

٩٥، ٨٦، ٨٧، ٩٠، ٦٢، ٧٣، ٨٩، ٩٢، ٨٤، ٧٦.

الحـل :

نرتب أولاً الدرجات السابقة ترتيبًا تصاعديًا مثلاً فتكون :

٦٢، ٧٣، ٧٦، ٨٤، ٨٦، ٨٧ ، ٨٩، ٩٠، ٩٢، ٩٥

ترتيب الوسيط الأول = $\frac{(ن)}{٢}$

= $\frac{١٠}{٢}$ = ٥ أى القيمة الخامسة فى الترتيب

∴قيمة الوسيط الأول = ٨٦

ترتيب الوسيط الثانى = $(\frac{ن}{٢}) + ١$

= $(\frac{١٠}{٢}) + ١$ = ٦ القيمة السادسة فى الترتيب

∴ قيمة الوسيط الثانى هى ٨٧.

$$الوسيط = \frac{الوسيط الأول + الوسيط الثانى}{٢}$$

$$الوسيط = \frac{٨٦ + ٨٧}{٢} = \frac{١٧٣}{٢} = ٨٦,٥$$

٢- حساب الوسيط من بيانات مبوبة (جداول تكرارية):

فـــى حالـــة البيانات المبوبة، يقوم الدارس بعمل الجدول المتجمع الصاعد أو الهابط وفق الترتيب. ثم يحدد بعد ذلك ترتيب الوسيط وذلك على النحو الآتى:

$$\text{ترتيب الوسيط} = \frac{\text{مجـ ك}}{٢}$$

والـــناتج مـــن العلاقـــة السابقة يحدد مباشرة الفئة التى يقع بين حديها الأدنى والأعلى ترتيب الوسيط. ويطلق على تلك الفئة (فئة الوسيط).

أ - حساب الوسيط باستخدام التكرار المتجمع الصاعد:

معادلة الوسيط فى حالة التكرار المتجمع الصاعد:

$$\text{الوسيط} = \text{الحد الأدنى لفئة الوسيط} + \frac{\text{ترتيب الوسيط} - \text{التكرار المتجمع السابق}}{\text{تكرار الأصلى للفئة}} \times \text{طول الفئة الوسيطية}$$

مثال: يوضـــح الجـــدول رقم (٤-٧) توزيع درجات ١٠٠ طالب فى امتحان مادة الفلسفة المعاصرة. أوجد الوسيط من هذه البيانات المبوبة.

خطوات الحل:

١- تكوين الجدول المتجمع الصاعد أو الهابط.

٢- حساب ترتيب الوسيط وهو يساوى $\frac{\text{مجـ ك}}{٢}$

٣- تطبيق معادلة قيمة الوسيط.

جدول رقم (٤-٧)
توزيع درجات الطلاب فى مادة الفلسفة المعاصرة

ف	ك
٤٠-	٥
٥٠-	٢٥
٦٠-	٣٥
٧٠-	٢٥
٨٠-٩٠	١٠
مجـ	١٠٠

حساب الوسيط بواسطة التكرار المتجمع الصاعد :

جدول رقم (٤-٩)
جدول التكرار المتجمع الصاعد

ف	ك	الحدود العليا	اتكرار المتجمع الصاعد
٤٠-	٥	أقل من ٥٠	٥
٥٠-	٢٥	أقل من ٦٠	٣٠
٦٠-	٣٥	أقل من ٧٠	٦٥ فئة الوسيط
٧٠-	٢٥	أقل من ٨٠	٩٠
٨٠-٩٠	١٠	أقل من ٩٠	١٠٠

$$\text{ترتيب الوسيط} = \frac{١٠٠}{٢} = ٥٠$$

$$\text{الوسيط} = ٦٠ + \frac{٥٠ - ٣٠}{٣٥} \times ١٠$$

$$= ٦٠ + \frac{٢٠٠}{٣٥}$$

$$= ٦٠ + ٥,٧١ = ٦٥,٧١ \text{ درجة}$$

ب- حساب الوسيط بواسطة التكرار المتجمع الهابط :

جدول رقم (٤-٩)
جدول التكرار المتجمع الهابط

ف	ك	الحدود الدنيا	تكرار متجمع هابط
٤٠-	٥	٤٠ فأكثر	١٠٠
٥٠-	٢٥	٥٠ فأكثر	٩٥
٦٠-	٣٥	٦٠ فأكثر	٧٠ الفئة الوسيطية
٧٠-	٢٥	٧٠ فأكثر	٣٥ (ك السابق)
٨٠-٩٠	١٠	٨٠ فأكثر	١٠

معادلة الوسيط فى حالة التكرار المتجمع الهابط هى:

$$\text{الوسيط} = \text{نهاية الفئة الوسيطة} - \frac{\text{ترتيب الوسيط} - \text{التكرار المتجمع السابق}}{\text{التكرار الأصلى للفئة الوسيطة}} \times \text{مدى الفئة}$$

$$\text{الوسيط} = ٧٠ - \frac{٥٠ - ٣٥}{٣٥} \times ١٠$$

$$= ٧٠ - \frac{١٥٠}{٣٥} = ٧٠ - ٤,٢٩$$

$$= ٦٥,٧١$$

وهى نفس النتيجة التى حصلنا عليها باستخدام جدول التكرار المتجمع الصاعد.

٣- استخدام منحنى المتجمع الصاعد والهابط فى إيجاد قيمة الوسيط:

يمكن باستخدام القيم الحقيقية أو النسب المئوية لتكرار الفئات إيجاد الوسيط بالرسم، إلا أن هذه الطريقة تعتبر أقل طرق حساب الوسيط دقة. وبنفس طريقة رسم المنحنى المتجمع الصاعد والهابط والتى أشرنا إليها بإيجاز فى الفصل السابق. يقوم الدارس برسم المنحنين الصاعد والهابط معًا لإيجاد الوسيط حتى يمكن تحقيق قدر مناسب من الدقة. ومن نقطة التقاء المنحنيين يتم إسقاط عمود على المحور الأفقى (السينى) (فئات) فيقطعه فى نقطة بعدها السينى يمثل قيمة الوسيط كما يتضح من الشكل رقم (٤-١).

مثال :

من بيانات جدول رقم (٤-٧) فى المثال السابق أوجد الوسيط باستخدام (الرسم) منحنى المتجمع الصاعد والهابط.

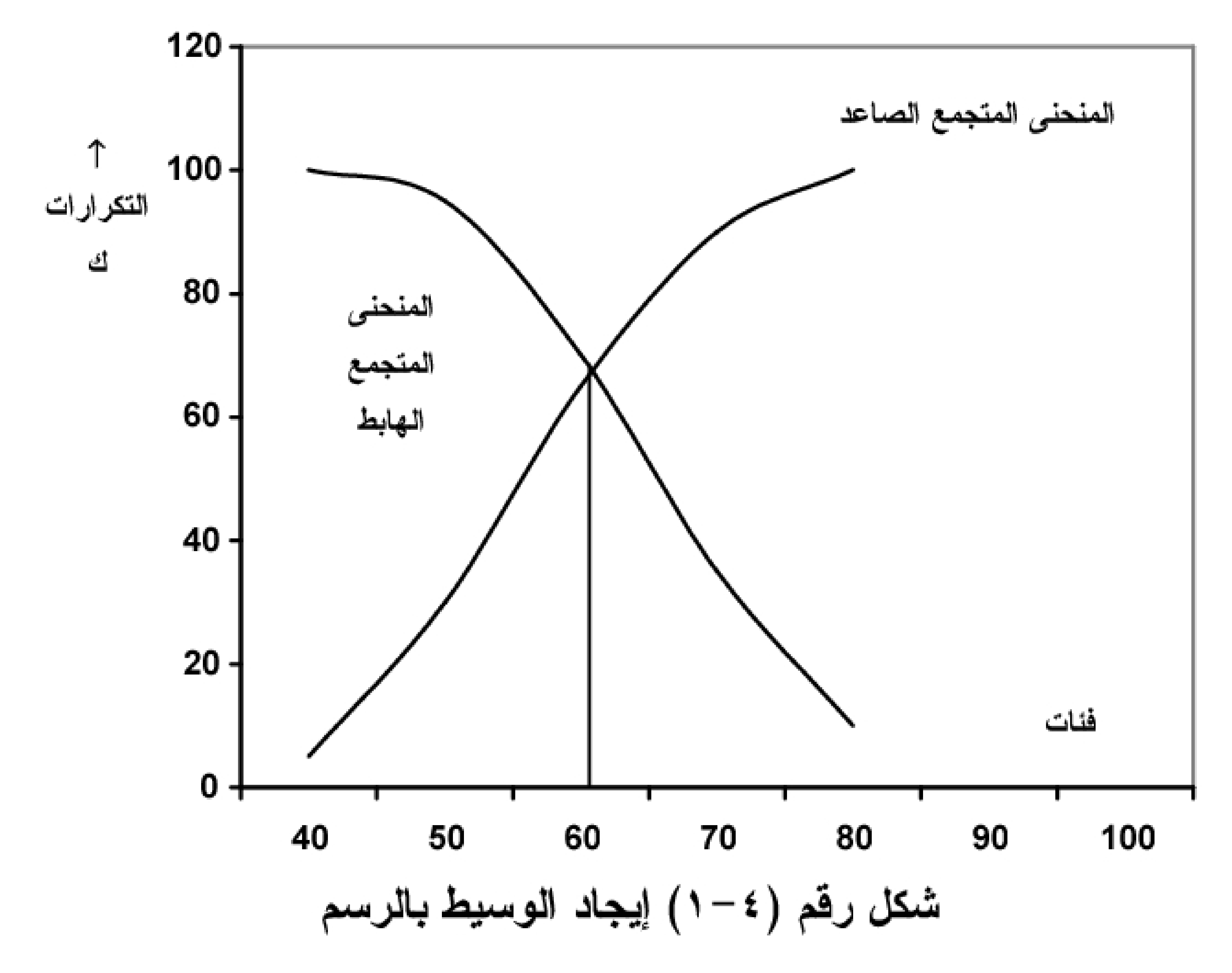
الحل :

عمل جدول تكرار متجمع صاعد وهابط.

رسم المنحنى المتجمع الصاعد والهابط بالطريقة التى سبق شرحها فى الفصل الثالث. وكما هو واضح فى الشكل رقم (٤-١).

جدول (٤-١٠)

ف	ك	الحدود العليا	تكرار متجمع صاعد	الحدود الدنيا	تكرار متجمع هابط
٤٠-	٥	أقل من ٥٠	٥	٤٠ فأكثر	١٠٠
٥٠-	٢٥	أقل من ٦٠	٣٠	٥٠ فأكثر	٩٥
٦٠-	٣٥	أقل من ٧٠	٦٥	٦٠ فأكثر	٧٠
٧٠-	٢٥	أقل من ٨٠	٩٠	٧٠ فأكثر	٥
٨٠-	١٠	أقل من ٩٠	١٠٠	٨٠ فأكثر	١٠
مجـ	١٠٠				



شكل رقم (٤-١) إيجاد الوسيط بالرسم

ثالثاً : المنوال :

يعـرف المـنوال Mode بأنه "القيمة أو الفئة الأكثر شيوعاً فى التوزيع" كما يعتـبر المنوال المقياس الوحيد من مقاييس النزعة المركزية الذى يستخدم لحساب المتوسـط (فى حالة البيانات الاسمية) مثل متغيرات المهنة،والجنس Sex، واللون والحالـة الزواجـية فكـل مـنها تمثل خاصية اسمية. والمنوال للبيانات المتقطعة

(Discrete) أو للبيانات الاسمية عبارة عن الفئة التى تحصل على أعلى التكرارات.

مثال :

الجدول التالى يوضح النسبة المئوية لتوزيع تكرارات لخمس فئات حرفية فى المجتمع الريفى. والمطلوب تحديد الفئة المنوالية بين تلك الفئات على ضوء التعريف السابق للمنوال.

جدول (٤-١١)

م	الفئات المهنية	%
١	مزارعون	٤٥
٢	تجار ماشية	٢٥
٣	حلاق صحة	١٥
٤	بناء	١٠
٥	حرف متنوعة	٥
		١٠٠

من التكرارات الموضحة بالجدول، يتضح أن فئة المزارعين تمثل الفئة المنوالية لاشتمالها على أعلى نسبة (٤٥%) هى الأعلى قياسًا بباقى التكرارات.

١ - حساب المنوال من القيم الخام Raw Data :

مثال :

فيما يلى درجات عدد من الطالبات فى مادة مبادئ الإحصاء.

٣٩ ، ٢٥ ، ١٤ ، ٣٩ ، ٢٥ ، ٢٠ ، ١٣ ، ١٦ ، ٢٥.

المطلوب: إيجاد قيمة المنوال.

خطوات الحل:

١- رتب القيم ترتيبًا تصاعديًا (الأصغر فالأكبر وهكذا ...).

٢- حدد القيمة أو القيم الأكثر شيوعًا.

١٣ ، ١٤ ، ١٦ ، ٢٠ ، ٢٥ ، ٢٥ ، ٢٥ ، ٣٩ ، ٣٩.

فى هذا المثال، نجد أن القيمة الأكثر شيوعًا هى (٢٥) وقد يكون هناك أكثر من منوال فى التوزيع.

نخلص مما سبق إلى تعريف عام للمنوال بأنه "القيمة الأكثر شيوعًا أو انتشارًا". ومن ثم يتوقف تحديد قيمة المنوال على تكرار القيم داخل المجموعة.

٢- طرق حساب المنوال من البيانات المبوبة (التوزيع التكرارى):

يمكن تقدير قيمة المنوال وتحديد الفئة المنوالية، إما بالطرق الحسابية باستخدام المعادلات وإما بطريقة الرسومات البيانية أو بكليهما معًا وسوف نعرض لخمس طرق لتقدير قيمة المنوال منها ثلاث حسابية وطريقتان بالرسم. والطرق الحسابية الثلاث هى :

أ - طريقة مركز الفئة المنوالية.
ب- طريقة استخدام التكرار السابق والتكرار اللاحق للفئة المنوالية.
ج- طريقة بيرسون (الفروق الدقيقة).

أ - تقدير المنوال باستخدام طريقة مركز الفئة المنوالية :

تعتبر هذه الطريقة أبسط الطرق الثلاث الحسابية وأقلهم دقة، نظرًا لأن المنوال عادة ما ينحاز إما صوب بداية الفئة المنوالية أو ناحية نهايتها تبعًا لتكرارات الفئتين السابقة واللاحقة للفئة المنوالية.

ومن ثم لا نتوقع تطابق قيمة المنوال مع مركز الفئة المنوالية إلا إذا تساوى تكرارى الفئتين السابقة واللاحقة للفئة المنوالية.

مثال:

احسب المنوال من التوزيع التكرارى للعمر لعينة من العاملين، وذلك باستخدام ثلاث طرق:

جدول (٤-١٢)

توزيع عينة العاملين حسب السن

فئة السن	ك
١٥-	١٥
٢٥-	٣٥ الفئة المنوالية
٣٥-	٢٥
٤٥-	١٥
٥٥-	٦
٦٥-	٤

أ - خطوات حساب المنوال باستخدام مركز الفئة:

١- تحديد الفئة المنوالية التى تشتمل على أعلى التكرارات والفئة هى (٢٥-٣٥).

٢- حساب مركز الفئة المنوالية:

$$\text{مركز الفئة} = \frac{\text{الحد الأدنى للفئة + الحد الأعلى}}{٢}$$

$$\therefore \text{مركز الفئة المنوالية} = \frac{٢٥ + ٣٥}{٢} = ٣٠$$

قيمة المنوال = ٣٠

ب- طريقة استخدام التكرار السابق والتكرار اللاحق للفئة المنوالية:

يمكـــن حساب المنوال باستخدام تكرارى الفئتين السابقة (ك١) واللاحقة (ك٢) للفئة المنوالية.

$$\text{المنوال} = \text{الحد الأدنى للفئة المنوالية} + \frac{\text{تكرار الفئة بعد المنوالية}}{\text{تكرار الفئة السابقة للمنوالية + تكرار الفئة اللاحقة}} \times \text{مدى الفئة المنوالية}$$

مثال :

احسب المنوال من التوزيع التكرارى لدخول عينة تتكون من ٢٥ عاملاً وذلك مـــن الجـــدول التالى باستخدام كل من طريقة تكرارى الفئتين (السابقة واللاحقة)، وطريقة بيرسون.

طريقة حساب المنوال باستخدام طريقة تكرارى الفئتين السابقة واللاحقة له.

جدول (٤-١٣)

فئات الدخل	ك
١٥-	٥
٢٠-	١٢
٢٥-	٤
٣٠-٣٥	٤
مجـــ	٢٥

وحيث إن طــول الفئـــة = ٥
الحد الأدنى للفئة المنوالية = ٢٠
تكرار الفئة قبل المنوالية = ٥
تكرار الفئة بعد المنوالية = ٤

$$\text{المنوال} = ٢٠ + \frac{٤}{٤ + ٥} \times ٥$$

= ٢٢,٢

ج- حساب المنوال باستخدام طريقة بيرسون (الفروق الدقيقة):

خطوات الحل :

أ - يــتم عمــل جــدول يضم الفئة المنوالية والفئتين السابقة واللاحقة لها بالفئات والتكرارات مع إضافة عمود للفروق.

ب- يحســب الفرق الأول بين تكرارى الفئة المنوالية والفئة السابقة لها، ثم يحسب الفرق الثانى بين تكرار الفئة المنوالية وتكرار الفئة اللاحقة لها.

تستخدم معادلة بيرسون لحساب قيمة المنوال وذلك على النحو التالى:

$$\text{المنوال} = \text{الحد الأدنى للفئة المنوالية} + \frac{\text{الفرق بين تكرارى المنوال والسابق}}{\text{الفرق بين تكرارى المنوال والسابق} + \text{الفرق بين تكرارى المنوال واللاحق}} \times \text{طول الفئة}$$

ويرمــز للفرق بين تكرارى المنوال والسابقة بالرمز (D_1) ويرمز للفرق بين تكــرارى المنوال واللاحق بالرمز (D_2) الحد الأدنى للفئة المنوالية بالحرف (ح د) وطول الفئة بالحرف (ط).

جدول (٤-١٤)

فئات	ك	فروق
١٥-٢٠	٥	
٢٠-٢٥	١٢	٧ D_1
٢٥-٣٠	٤	٨ D_2

$$\therefore \text{المنوال} = \text{ح} + \frac{D_1}{D_1 + D_2} \times \text{ط}$$

$$\text{المنوال} = ٢٠ + \frac{٧}{٧ + ٨} \times ٥$$

$$= ٢٠ + ٢,٣ = ٢٢,٣$$

وهى تقريبًا نفس القيمة السابقة مع قدر يسير من الدقة.

ولعــل الاختلاف البسيط بين قيمتى المنوال يدل على أن المنوال مقياس غير مستقر نجد أن قيمته تتوقف على تبويب البيانات فى حالة التوزيعات التكرارية، فلو كان التبويب متصفاً باختلاف أطوال الفئات لاختلفت تبعًا لذلك قيمة المنوال.

ففى المثال السابق راعينا أن تكون أطوال الفئات متساوية، ولكن هناك حالات لا تــتحقق فــيها هذه الخاصية. وفى مثل تلك الحالات، لا يستخدم الباحث معادلة بيرسون أو أى من المعادلات السابقة.

٣- حساب المنوال إذا كانت أطوال الفئات غير متساوية:

فــى هــذه الحالة يقوم الباحث بإجراء تعديلات فى التكرارات المدونة بجدول البيانات وذلك بأن يقسم كل تكرار على طول الفئة المناظرة لهذا التكرار. وفى هذه الحالة يمكن للباحث أن يستخدم معادلة بيرسون السابقة.

أمـــا إذا أراد الباحث أن يستخدم التكرارات الأصلية دون إدخال أى تعديلات فيمكن أن يحسب المنوال فى هذه الحالة من المعادلة التالية :

$$\text{المنوال} = \text{الحد الأدنى للفئة المنوالية} + \frac{\text{طول الفئة قبل المنوالية} + \text{تكرار الفئة المنوالية}}{\text{طول الفئة بعد المنوالية} \times \text{تكرار الفئة قبل المنوالية} + \text{طول الفئة قبل المنوالية} \times \text{تكرار الفئة بعد المنوالية}} \times \text{مدى الفئة بعد المنوالية}$$

٤- حساب المنوال من الرسومات البيانية :

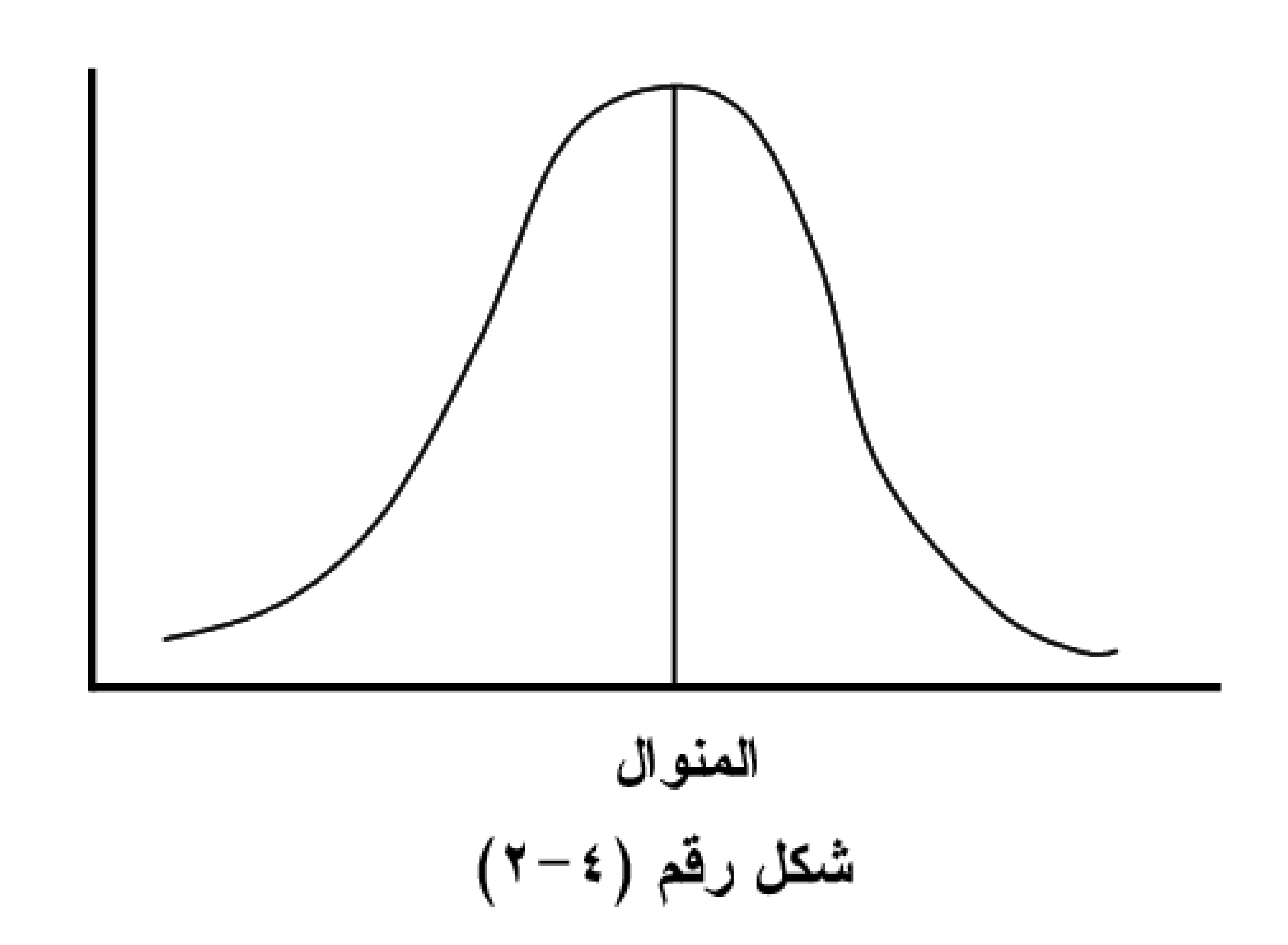
أ - من المنحنى التكرارى Frequency Curve :

مــن التعريف العام للمنوال بأنه دائمًا القيمة التى تقابل أكبر تكرار. نقول أننا إذا عرضـــنا التوزيع التكرارى باستخدام المنحنى التكرارى وقمنا بإسقاط عمود من

أعلـــى نقطة فى المنحنى (التى تمثل أعلى قيمة تكرارية) فإنه سوف يقطع المحور الأفقى أو السينى فى نقطة هى المنوال. كما يتضح ذلك من الشكل رقم (٤-٢).

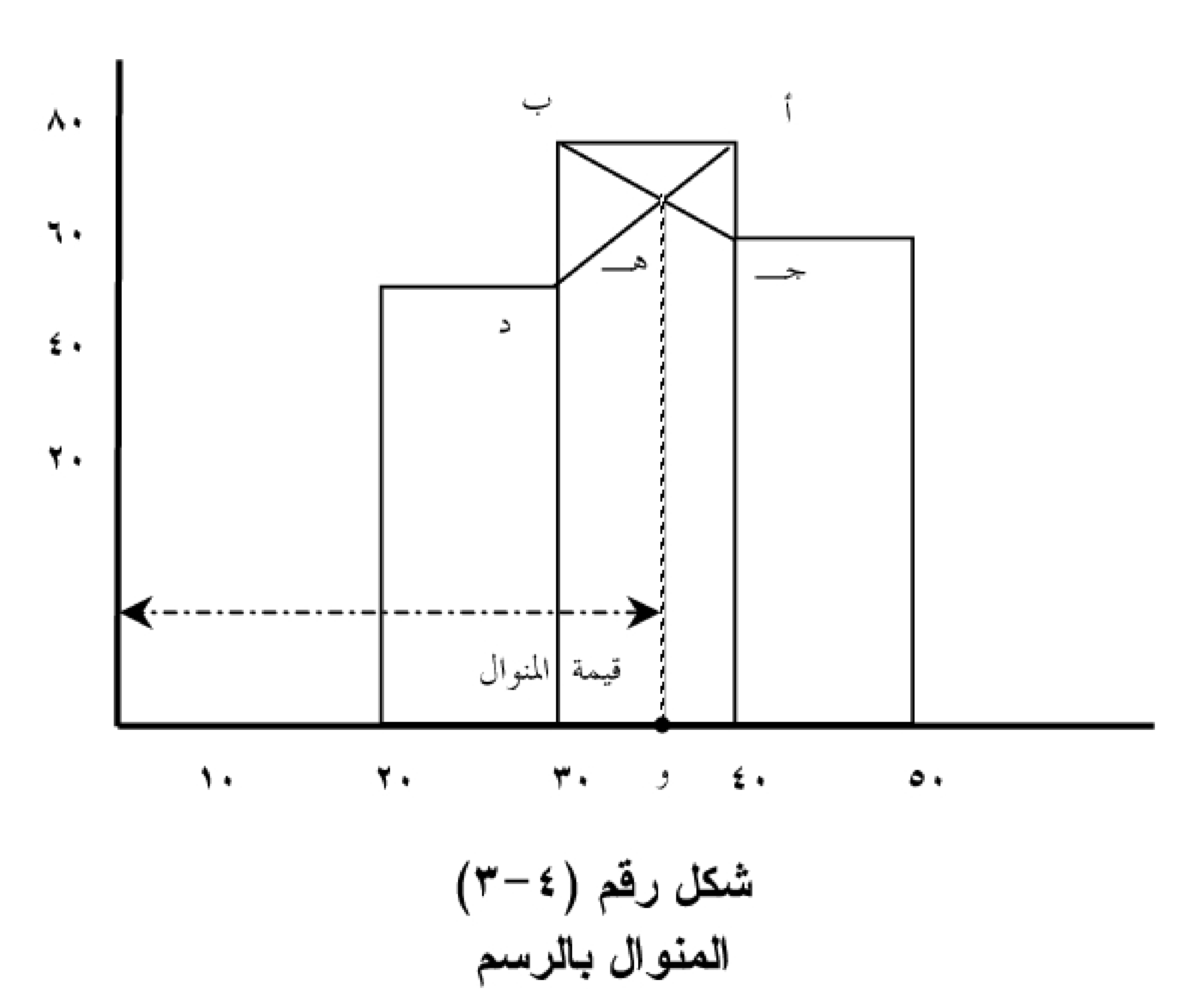
ب- من المدرج التكرارى:

يكـــتفى عند رسم المدرج التكرارى لحساب المنوال باختيار ثلاثة مستطيلات فقط يمثل الأوسط الفئة المنوالية وعلى كل جانب منه يقوم الدارس برسم: مستطيل



المنوال

شكل رقم (٤-٢)



شكل رقم (٤-٣)
المنوال بالرسم

يمـــثل القـــيم التكرارية للفئة السابقة مباشرة للفئة المنوالية ويكرر نفس العمل على الجهـــة الأخـــرى مـــن مستطيل الفئة المنوالية حيث يقوم برسم مستطيل يمثل القيم الـــتكرارية للفئة التالية مباشرة للفئة المنوالية. كما يتم رسم خط محورى يصل من بداية الفئة المنوالية ببداية الفئة اللاحقة وليكن (ب، جـ) وخط محورى آخر يصل نهاية الفئة المنوالية مع نهاية الفئة السابقة (أ ، د).

لـــو افترضـــنا أن طـــرفى الفئة المنوالية هما (أ ، ب) كما فى الشكل رقم (٤-٣) يرسم الدارس خط محورى ليصل بداية الفئة المنوالية ليصل ببداية الفئة اللاحقـــة (ب، جـــ) ويرسم خط محورى آخر من نهاية الفئة المنوالية ليصل بنهاية الفئة السابقة وليكن (هـ) (أ ، د) ومن نقطة تلاقى الخطين يتم رسم خط عمـــودى يقطـــع المحـــور السينى فى نقطة ولتكن (و) وهذه النقطة تمثل قيمة المنوال.

رابعاً: ملاحظات على مقاييس النزعة المركزية :

المقياس	المزايا	العيوب
أولاً: المتوسط الحسابى	١- يعتبر أفضل المقاييس الثلاثة لـــتقدير الوســـط الحســـابى الحقيقى للمجتمع الأصلى. ٢- يتميز عن الوسيط والمنوال بأنـــه يستخدم جميع البيانات المـــتاحة عـــن الظاهـــرة موضوع الدراسة. ٣- يعتبر أفضل المقاييس الثلاثة فى حالة القياسات الكمية.	١- يصعب تقدير المتوسط الحسابى بدقـــة من التوزيعات التكرارية المفتوحة. ٢- استبعاد كل الفئات المفتوحة فى حساب المتوسط

المقياس	المزايا	العيوب
ثانيًا : الوسيط	١- أفضل المقاييس الثلاثة إذا كانت القياسات التى تسجل عن الظاهرة ترتيبية. ٢- يفضل عن المتوسط الحسابى كلما إزدادت درجة التواء التوزيعات التكرارية، نظرًا لأن الوسيط لا يتأثر غالبًا بالقيم المتطرفة للظاهرة. ٣- يتميز بإمكانية استخدامه حتى فى حالة عدم معرفة القيم الكبرى أو الصغرى فى جداول التوزيعات.	١- صعوبة استخدامه فى عمليات جبرية. ٢- لا يمكن حساب الوسيط العام لعدة مجموعات من البيانات.
ثالثًا : المنوال	١- يعتبر افضل المقاييس بل المقياس الوحيد للمتوسط فى حالة البيانات الاسمية. ٢- يعتبر أفضل المقاييس شيوعًا للتعبير عن شكل وتوزيع البيانات. ٣- يمتاز بسهولة حسابه. ٤- لا يتأثر بالقيم الشاذة. ٥- يمكن حسابه فى حالة التوزيعات المفتوحة خاصة البيانات الاسمية فيعتبر أفضل المتوسطات الثلاثة تمثيلاً لتلك البيانات.	١- يتأثر بتغيير أطوال الفئات مما يقلل من أهميته ويحد من استخداماته. ٢- مقياس غير مستقر تتوقف قيمته فى حالة التوزيعات التكرارية على طريقة التبويب.

العلاقة بين المتوسطات الثلاثة للنزعة المركزية :

نستنتج من دراستنا للمقاييس الثلاثة للنزعة المركزية أن التوزيع يكون متماثلاً مــثل المنحــنى الاعــتدالى (الجرسى) إذا تساوت قيم المتوسط الحسابى والوسيط والمــنوال وعندما تكون قيمة المتوسط الحسابى أصغر من قيمة المنوال بفارق يقل عــن الصفر يكــون التوزيع سالب الالتواء Negative skewness. أمــا إذا كان المتوســط الحســابى أكبر من قيمة المنوال بفارق موجب أى أكبر من الصفر فإن الــتوزيع التكرارى يكون موجب الالتواء Positive skewness وقد أمكن باستخدام بعــض العملــيات الرياضية البسيطة إيجاد معادلة تربط بين المقاييس الثلاثة على النحو الآتى :

المنوال = المتوسط الحسابى – ٣ (المتوسط الحسابى – الوسيط)

خامساً : المفاهيم الأساسية Key Concepts

١ - المتوسط الحسابى Arithmetic Mean :

هو أكثر مقاييس النزعة المركزية استخدامًا ونحصل عليه بقسمة مجموع القيم على عددها.

٢ - الوسيط Mdeian:

تعـــنى كلمة الوسيط منتصف الشيء. فهو القيمة التى تقع فى المنتصف تمامًا بعـــد ترتيب القيم تنازليًا أو تصاعديًا. أى القيمة التى يكون عدد القيم الأخرى الأقل منها مساويًا تمامًا لعدد القيم الأعلى منها. والوسيط أفضل مقاييس النزعة المركزية استخدامًا فى حالة التواء التوزيع.

٣ - المنوال Mode:

المنوال لمجموعة من القيم هو القيمة التى تقابل أكبر تكرار فى المجموعة أو بمعنى آخر هى القيمة الأكثر شيوعًا.

سادساً : تمـــارين

١- فـــيما يلى جدول توزيع تكرارى لـــ (٥٠٠) عاملاً بأحد المصانع، والمطلوب قـــياس متوســـط العمـــر لهـــم باستخدام المنوال مع بيان أفضل المتوسطات استخدامًا فى تمثيل أعمارهم وسبب هذا التفضيل.

الأعمار	عدد العمال
١٥-	٤٥
٢٥-	٢٣٥
٣٥-	١٥٠
٤٥-	٥٠
٥٥-٦٥	٢٠
	٥٠٠

٢- احسب كل من الوسيط والمتوسط الحسابى والمنوال للأجور التالية. ثم بين أى المقاييس أفضل.

٣,١٠	٢,٥٧	٢,٥٠	٧,٨٠	٣,٧٥
٣,٢٨	٣,٩٦	٢,٥٧	٣,٦٠	٢,٥٧

٣- الجـــدول الآتى يبين عدد المدمنين حسب فئاتهم العمرية فى إحدى المدن عام ٢٠٠٩.

الفئات	عدد الحالات
١٥-	٧٠
٢٠-	١٢٠
٢٥-	٤٦٠
٣٠-	٢٠٠
٣٥-	١٦٧
٤٠-	٨٢
٤٥-	٢٧
٥٠-٥٥	١٥

والمطلوب:

أ - عمل جدول تكرارى مئوى.

ب- رســـم المـــدرج التكرارى لهذا التوزيع وأوجد منه قيمة المنوال ثم حقق الناتج حسابيًا.

٤- فــيما يلـــى توزيـــع تكرارى لبيانات افتراضية عن الدخل السنوى لعينة من الأسر، احسب المنوال:

فئات الدخل بالجنيه	ك
-٧٠	١٦
-٩٠	١١٨
-١١٠	٤٠٦
-١٣٠	١٢٦٤
-١٥٠	١٦١١
-١٧٠	١١٥١
-١٩٠	٤٩٧
-٢١٠	٢٣٨
-٢٣٠	٧٠
٢٥٠-٢٧٠	٣٢

٥- فــيما يلــى درجــات خمســين طالــبًا فى الامتحان النهائى لمادة الإحصاء الاجــتماعى، والمطلــوب جدولـــة هذه البيانات فى توزيع تكرارى مئوى ثم احسب المتوسط الحسابى والمنوال من هذا الجدول:

١٤	١٥	١٦	١٩	١٧
١٧	١٨	١٤	١٤	٢٠
١٢	١٤	١٢	١٩	١٤
١٠	١٢	١٥	٢٠	٨
١٣	١١	١٩	١٥	٩
٨	١٢	٢٠	١٦	١٠
١٤	١٤	١٧	١٧	٢٠
١٠	١١	٢٠	١٣	١٥
١٥	١٥	١٠	٨	١٣
١٢	١٢	١٣	٩	١٧

٦- فيما يلى مجموعة من القيم والمطلوب حساب المتوسطات الثلاثة:

٦ ، ٢ ، ٤ ، ٧ ، ٢٠ ، ٤ ، ٧ ، ٨ ، ٨ ، ٩ ، ١٢ ، ٧ ، ٥ ، ٦ ، ١٩ ، ٢٠ ، ٦ ، ٨ ، ١١ ، ٧ ، ٩.

٧- فيما يلى درجات ٣٠ طالبًا فى امتحان مادة مناهج البحث :

٦٥	٧٤	٣٥	٣٨	٧٠
٩٥	٥٨	٧٤	٧٠	٦٥
٧٣	٣٥	٧٠	٧٠	٦٧
٤٦	٧٠	٤٨	٣٥	٦٦
٣٠	٩٦	٣٥	٩٦	٤٨
٩٥	٤٨	٩٥	٧٠	٩٦

المطلــوب جدولـــة هذه الدرجات الخام وحساب الوسيط والمتوسط الحسابى. على افتراض أن المتغير متصل.

٨- تمــثل البيانات الآتية أوزان افتراضية لعينة من المبحوثين والمطلوب جدولة تلك الأوزان وحساب الوسيط.

٨٠	٨٥	٨٤	٦٠	٦٥
٤٠	٣٢	٨١	٦٠	٥٣
٩٥	٩٥	٨٠	٩٥	٨٤
٦٣	٨٠	٧٣	٨٠	٩١
٧٠	٥٢	٦٣	٤٠	٦٣

٩- احســب المتوســط الحسابى والوسيط من الجدول التكرارى الآتى مع توضيح أنســب مقايــيس النزعة المركزية فى حالة وجود التواء فى توزيع القيم، ثم وضح كلاً من المنحنى المتجمع الصاعد والهابط بالرسم:

الفئات	عدد الحالات
١٠-	٢
٢٠-	١٩
٣٠-	٩
٤٠-	٣
٥٠-	٢١
٦٠-	٧
٧٠-	٨
٨٠-	١٥
٩٠-١٠٠	٦

١٠- احسب كلاً من الوسيط والمنوال من بيانات جدول التوزيع التكرارى التالى:

الفئات	عدد الحالات
٢٠-	١
٢٥-	٢
٣٠-	٥
٣٥-	٢٠
٤٠-	٢٢
٤٥-	٤٢
٥٠-	٣٠
٥٥-	١٠
٦٠-	١٥
٦٥-٧٠	٦

١١- أكمل ما يأتى :

(أ) تشير عبارة النزعة المركزية إلى .... حالة فى وجود التوزيع.

(ب) المقاييس الثلاثة الأساسية للنزعة المركزية هى المنوال، ....، ....

(ج) إن الغرض من جميع مقاييس النزعة المركزية هو .... للتوزيع الكلى للقيم بواسطة وصف أكثر القيم المتطابقة داخله.

(د) على النقيض من المنوال، فإن .... يمثل دائمًا المركز الحقيقى لتوزيع القيم.

(هـ) إن .... يعتبر أكثر المقاييس استخدامًا للنزعة المركزية إلا أنه يتطلب مراجعة كاملة فقط فى حالة استخدامه مع مستوى بيانات ....

(و) يتأثر مقياس .... بكل قيمة داخل التوزيع.

١٢-اختر الإجابة الصحيحة من بين الاختيارات التى تلى كل سؤال وذلك بوضع دائرة على الاختيار الصحيح فيما يلى:

- **إن السبب الرئيسى لحساب مقاييس النزعة المركزية يتمثل فى:**

(أ) تلخيص المتغيرات الفردية.

(ب) إيجاد قيمة متوسطة.

(ج) معرفة خصائص المتغير.

(د) لا إجابة صحيحة من الإجابات الثلاث السابقة.

(هـ) جميع الإجابات الثلاث السابقة.

- **يعرف الوسيط بأنه النقطة التى:**

(أ) تمثل أعلى التكرارات المشاهدة.

(ب) تمثل القيمة التى عندما يتم طرحها من المتوسط يكون المنوال هو النتيجة.

(ج) تمثل المحل المركزى داخل التوزيع.

(د) تتمثل فى جميع الإجابات الثلاث السابقة.

(هـ) لا تتمثل فى أى من الإجابات السابقة.

- **إن أكثر المتغيرات استخدامًا، فى البحث الاجتماعى هو الرقم المتوسط لسنوات التعليم التى يتلقاها المصريون فى مراحل التعليم المختلفة. وهذا الرقم هو :**

(أ) المنوال.

(ب) الوسيط.

(ج) المتوسط.

(د) ليس واحدًا من المقاييس السابقة.

(هـ) واحد فقط من المقاييس السابقة.

- **لو أن أستاذ مادة النظرية الاجتماعية قد أعلن أن كل درجة حصلت عليها طالبات السنة الثالثة من قسم الاجتماع قد زادت سبع درجات لأنه استبعد سؤالين من أسئلة الامتحان وأضيفت درجاتهما إلى باقى أسئلة الامتحان ففى هذه الحالة ماذا حدث للمتوسط الجديد؟**

(أ) لم يتغير.

(ب) يساوى المتوسط القديم مع إضافة قيمة تساوى ($\frac{٧}{ن}$)

(ج) يساوى المتوسط القديم مضافاً إليه (٧).

(د) يساوى المتوسط القديم مضافاً إليه قيمة مقدارها $\frac{١٤}{٢}$ درجة.

(هـ) لا توجد معلومات كافية للإجابة على السؤال.

- **فى إحدى التوزيعات النوعية، كانت قيمة المنوال (٧٥)، والوسيط (٧٠)، والمتوسط (٦٥). فهل يكون هذا التوزيع :**

(أ) اعتياديًا Normal.

(ب) يتصف بالالتواء الموجب.
(ج) يتصف بالالتواء السالب.
(د) متماثل Symmetrical.
(هـ) لا يتصف بأى صفة من الصفات الأربع السابقة.

- **يستخدم المنوال إحصائيًا فى قياس النزعة المركزية عندما تكون القياسات المعطاة :**

(أ) اسمية Nominal.
(ب) ترتيبية.
(ج) فاصلة.
(د) نسبة.
(هـ) جميع الخصائص السابقة.

- **لو كان التوزيع المعطى متصفًا بالتماثل، فإن أفضل مقاييس النزعة المركزية استخدامًا، فى هذه الحالة هو :**

(أ) المنوال.
(ب) الوسيط.
(ج) المتوسط.
(د) جميع المقاييس المذكورة سابقًا.

١٣- احسب المنوال من الحسابات الآتية :

١- المتوسط الحسابى = ٤٦. الوسيط = ٤٠.
٢- المتوسط الحسابى = ٣٥. الوسيط = ٣٢.
٣- المتوسط الحسابى = ٢١٠. الوسيط = ٢٠٥.

٥

الفصل الخامس
مقاييس التشتت

الفصل الخامس
مقاييس التشتت

مقدمة :

تناولـــنا فى الفصلين الثالث والرابع خاصيتين أساسيتين من خصائص التوزيع هما الشكل والتعبير عنه بالرسومات البيانية، ثم مقاييس النزعة المركزية. ومن ثم تتــبقى خاصـــية ثالـــثة هـــى التشتت أو درجة تباين القيم داخل التوزيع موضوع الدراســة. فإذا كانت مقاييس النزعة المركزية تمد الباحث بقيمة واحدة تصف حالة الـــتوزيع للقـــيم جملـــة واحدة، فإن للتشتت أهمية فى قياس الفروق الفردية داخل التوزيع، بمعنى آخر توضح مقاييس التشتت (التباين) Measures of Dispersion (variation) درجـــة الـــتقارب أو التباعد بين القيم فى العينة موضوع الدراسة عن وسطها الحسابى. وتوجد مقاييس عديدة متاحة فى الإحصاء الوصفى لقياس التباين للمتغير المتصل هى:

١- المدى The Range.
٢- الانحراف الربيعى The quartile Deviation.
٣- الانحراف المتوسط The mean deviation.
٤- التباين The Variance.

تمثل مقاييس التشتت مؤشرات دالة على وقوع اختلافات فى توزيع ظاهرة ما موضـــوع الدراسة. وجدير بالذكر أن مقاييس النزعة المركزية والأشكال البيانية لا تكفـــى لوصف توزيع ظاهرة ما أو لعقد مقارنات بين مجموعة وآخرى. وفيما يلى مثالاً يوضح أنه رغم تساوى قيم المتوسط الحسابى لمجموعتين فهذا لا يعنى وجود اتساق فى القيم حول المتوسط، حيث إن واقع التوزيع يشير إلى اختلاف بينهما من حيث تشتت المفردات.

مثال:

أجـــرى باحـــث اجتماعى دراسة حول التماسك الأسرى لعدد من أسر الأطباء والمدرسين وكانت النتائج على النحو التالى :

الأطباء	المدرسون
١٣	٤
١٥	٣٧
١٣	٢٩
١٤	٧
١٧	١٠
١٤	٣٤
١٦	٥
١٨	٩
١٩	٨
١١	٧
مجـ ١٥٠	١٥٠

$\overline{س}$ الأطبـــاء = $\frac{١٥٠}{١٠}$ = ١٥

$\overline{س}$ المدرسين = $\frac{١٥٠}{١٠}$ = ١٥

يتضـــح تســـاوى قيمة الوسط الحسابى للمجموعتين. ولكن إذا تفحصنا توزيع الدرجـــات حول وسطها الحسابى أى مدى قربها أو بعدها منها لوجدنا توزيع القيم أكثر تشتتا فى مجموعة المدرسين عنه فى مجموعة الأطباء. فإذا استخدمنا الوسط الحســـابى فقط لعقد المقارنة بين المجموعتين لكان مضللاً. ومن ثم كان ضروريًا يقترن المتوسط بمعامل آخر هو التشتت.

مـــثال: وضـــح باســـتخدام مقياس المدى أى التوزيعين الآتيين أكثر تشتتًا من الآخر:

التوزيع الأول	١١	١٦	١٨	٢٣	٢٩	٣١	٣٧
التوزيع الثانى	١٨	١٩	٢١	٢٣	٢٤	٢٦	٢٩

الحـــل:

المدى للتوزيع الأول = (٣٧ – ١١) + ١ = ٢٧.

المدى للتوزيع الثانى = (٢٩ – ١٨) + ١ = ١٢.

أى أن الـــتوزيع الأول أكـــثر تشـــتتا فـــى قيمه من التوزيع الثانى علمًا بأن التوزيعين لهما قيمة متماثلة للوسيط (٢٣).

ومـــن عـــيوب استخدام المدى عدم صلاحيته للتطبيق على المجتمع الأصلى والعيـــنات كبيرة الحجم، بينما يسهل حسابه على العينات صغيرة الحجم حيث تكون الفرصة أفضل لاشتمالها على قيم شاذة عليا ودنيا. ومن ثم نادرًا ما يستخدم علماء الاجـــتماع هذا المقياس إلا فى الدراسات الكشفية أو الاستطلاعية فى معظم الأحيان (Hinkle et al 1979 : 43 & Kurtz, 1983 71-72).

حساب المدى فى جدول تكرارى :

يمكن حساب المدى للبيانات المبوبة بطريقتين كما سبق وأن أوضحنا:

الطريقة الأولى :

المدى المطلق: الحد الأعلى لأعلى فئة – الحد الأدنى لأدنى فئة.

الطريقة الثانية :

المدى المطلق: مركز الفئة الأعلى – مركز الفئة الأدنى.

ومن ثم يشتمل الفصل على العناصر الآتية:

١- مقاييس التباين للمتغير المتصل.
٢- مقياس التباين للمتغير المتقطع.
٣- المفاهيم الأساسية.
٤- تمارين.

فى هذا السياق، يهدف هذا الفصل إلى أن :

١- أن يعـــرف الدارس معنى مقاييس التشتت وكيفية حسابها للمتغيرات المتصلة والمتغيرات المتقطعة، وأن يعرف نقاط القوة والضعف لكل مقياس.
٢- يحدد كل من مقاييس النزعة المركزية ومقاييس التشتت الملائمة لكل مستوى من مستويات القياس (الاسمى، الرتبى، الفاصلة والنسبة).

أولاً : مقاييس التباين للمتغير المتصل :

١- المدى

يعتـــبر المدى Range أبسط مقاييس التشتت وأسهلها فى الحسابات، وله مزايا كمـــا أن له عـــيوبًا. فمن مزاياه أنه يكون نافعًا فى الحالات التى تتطلب سرعة فى الحســـابات والحصول على مؤشرات أولية عن التشتت لتوزيع ما. كما يكون مفيدًا للمبتدئيـــن الذين لا تتوفر لديهم المهارة الإحصائية الكافية لاستخدام مقاييس تشتت أكثر تعقيدًا.

ويكون استخدام مقياس المدى كذلك أكثر نفعًا فى حالة البيانات التى لا تتطلب معالجـــة إحصـــائية متقدمة أو البيانات التى سبق دراسة خصائصها وتتطلب مجرد الإلمام بما تتصف به من تشتت.

كيفية استخدام المدى فى قياس التشتت:

يعـــرف المـــدى بأنـــه الفرق الحسابى المطلق (أى بدون استخدام الإشارات الموجـــبة أو الســـالبة للقيم) بين أعلى قيمة وأقل قيمة فى التوزيع فى حالة البيانات غـــير المـــبوبة. أو الفرق بين الحد الأعلى لأعلى فئة والحد الأدنى لأدنى فئة فى الـــتوزيعات التكرارية. كما يعرف بأنه الفرق بين مركز الفئة الأعلى ومركز الفئة الأدنى.

أ - حساب المدى للقيم غير المبوبة:

مثال: فيما يلى أوزان عدد من الأطفال والمطلوب حساب المدى المطلق.

١٣ ، ١٥ ، ٣٥ ، ١٧ ، ٩ ، ٢٠ ، ١٨ ، ١٦

خطوات الحل:

١- ترتيب القيم إما تصاعديًا أو تنازليًا.
٢- تحديد أكبر قيمة وأصغر قيمة فى التوزيع.
٣- طرح أصغر قيمة من أكبر قيمة.

ترتيب القيم:

٩ ١٣ ١٥ ١٦ ١٧ ١٨ ٢٠ ٣٥

∴ المدى المطلق = ٣٥ – ٩ = ٢٦ + ١ = ٢٧ كيلو جرام.

أما عن عيوبه فمن عيوب المدى أنه يعتمد فى حساب التشتت على أعلى قيمة وأصغر قيمة فى التوزيع. وقد يصعب الحصول عليهما فى العينة البحثية. ومن ثم

تلعب الصدفة دورًا فى تكرار حدوث أعلى وأقل القيم داخل توزيع بعينه. فإذا افترضنا أن باحثًا أراد أن يعرف التشتت فى توزيع الثروة داخل مجتمع محلى، ولا يوجد بداخله سوى مليونير واحد فقط. فلو اختار الباحث عينة من عشرة أو عشرين فردًا من هذا المجتمع، فإن احتمالات اشتمال هذه العينة على المليونير الوحيد ستكون ضعيفة جدًا. ومن ثم تكون قيمة المدى الدالة على التشتت فى توزيع الثروة مضللة (Graham, 1994: 80 & Blalock 1972 : 78).

ومن عيوب المدى أيضا، أن الباحث عندما يستخدمه لا يستطيع أن يتعرف على درجة التباين للقيم الواقعة بين أعلى القيم وأدناها فى التوزيع. لهذا السبب نجد أن الباحثين فى استخدامهم للمدى يضيفون إلى الفرق بين أعلى قيمة وأقل قيمة مقدار الواحد الصحيح حتى تغطى قيمة المدى الدالة على تشتت التوزيع أعلى القيم وأقل القيم. ويتم حساب المدى باستخدام المعادلة الآتية :

المدى = (أعلى قيمة – أقل قيمة) + ١

ب- إيجاد المدى للجداول التكرارية :

مثال:

توضح بيانات الجدول الآتى توزيع عينه بحثية من العاملين فى إحدى الشركات الصناعية حسب فئات الأجور الأسبوعية.

المطلوب: ايجاد قيمة المدى المطلق.

فئات الأجور	العدد
١٠ –	٥
٢٠ –	٥
٣٠ –	١٣
٤٠ –	١٤
٥٠ –	١٦
٦٠ –	٨
٧٠ –	٦
٨٠ – ٩٠	٣
مجـ	٧٠

١- إيجاد قيمة المدى المطلق باستخدام الطريقة الأولى:

الحد الأدنى لأدنى فئة = ١٠

الحد الأعلى لأعلى فئة = ٩٠

∴ المدى المطلق = ٩٠ – ١٠ = ٨٠

٢- إيجاد قيمة المدى المطلق باستخدام الطريقة الثانية :

مركز الفئة الأدنى = ١٥

مركز الفئة لأعلى = ٨٥

∴ المدى المطلق = ٨٥ – ١٥ = ٧٠

٢- الانحراف الربيعى :

نظرًا لأن المدى يعتمد فى حسابه على القيمتين : الأعلى والأدنى فى التوزيع، فضلاً عن إعتماده على عدد الحالات، فإنه يعتبر مقياسًا غير مستقر. وهذا القصور فى المدى يمكن التغلب عليه بإيجاد الانحراف الربيعى. ويعرف بنصف المسافة بين الربيعين الأول والثالث. فلو رمزنا للانحراف الربيعى بالرمز (ر) وللربيع الأول بالرمز (ر١)، وللربيع الثالث (ر٣) يمكن حساب الانحراف الربيعى من المعادلة التالية :

$$\text{الانحراف الربيعى} = \frac{\text{ر٣} - \text{ر١}}{٢}$$

ويستخدم مقياس نصف المدى الربيعى أو الانحراف الربيعى عادة فى المستوى الرتبى للبيانات. وهذا المقياس أكثر استخدامًا فى البحوث التربوية والنفسية ويندر استخدامه فى البحوث الاجتماعية (Graham, 1994 : 80, 81) وفيما يلى خطوات حساب الانحراف الربيعى.

١- حساب الانحراف الربيعى من البيانات غير المبوبة :

مثال: فيما يلى درجات عدد من الطالبات فى امتحان الإحصاء الاجتماعى والمطلوب حساب درجة التشتت باستخدام مقياس الانحراف الربيعى.

١١	٩	١٥	١٧	٨	٧
١٤	١٣	٢٠	١٨	١٦	٥

خطوات الحل :

١- ترتيب القيم تصاعديًا أو تنازليًا
٢- تحديد قيمة الربيع الأدنى ($ر_١$)
٣- تحديد قيمة الربيع الأعلى ($ر_٣$)

تطبيق المعادلة $\frac{ر_٣ - ر_١}{٢}$

ترتيب القيم :

٥ ٧ ٨ ٩ ١١ ١٣ ١٤ ١٥ ١٦ ١٧ ١٨ ٢٠

قيمـــة $ر_١$ = $\frac{عدد القيم}{٤} = \frac{ن}{٤}$

= $\frac{١٢}{٤} = ٣$

∴ قيمة $ر_١$ = القيمة الثالثة وهى حسب الترتيب ٨

قيمـــة $ر_٣$ = $\frac{ن}{٤} \times ٣$

= $\frac{١٢}{٤} \times ٣$

∴ قيمة $ر_٣$ هى القيمة التاسعة وقيمتها ١٦

الانحراف الربيعى = $\frac{ر_٣ - ر_١}{٢}$

= $\frac{١٦ - ٨}{٢} = ٤$

٢- حساب الانحراف الربيعى من الجداول التكرارية:

مثال: يتضـــمن الجدول التالى توزيع تكرارى لأوزان عينة من الأطفال والمطلوب حساب قيمة نصف المدى الربيعى.

الوزن	٤ –	٨ –	١٢ –	١٦ –	٢٠ –	٢٤ –	٢٨–٣٢
العدد	٣	٥	٦	٤	١٠	٧	٥

خطوات الحل :

١- عمل جدول متجمع صاعد للبيانات.

٢- تحديد موقع الربيع الأدنى = $\frac{\text{مجـ ك}}{٤}$

٣- تحديد موقع الربيع الأعلى = $\frac{\text{مجـ ك}}{٤} \times ٣$

ف	ك	الحدود العليا	تكرار متجمع صاعد
٤ –	٣	أقل من ٨	٣
٨ –	٥	أقل من ١٢	٨
١٢ –	٦	أقل من ١٦	١٤ موقع $ر_١$
١٦ –	٤	أقل من ٢٠	١٨
٢٠ –	١٠	أقل من ٢٤	٢٨
٢٤ –	٧	أقل من ٢٨	٣٥ موقع $ر_٣$
٢٨–٣٢	٥	أقل من ٣٢	٤٠
مجـ	٤٠		

موقع الربيع الأدنى ($ر_١$) = $\frac{٤٠}{٤}$ = ١٠

قيمة $ر_١$ = $ح_د$ + $\frac{\text{موقع } ر_١ - \text{ك سابق}}{\text{ك الفئة الأصلية } ر_١}$ × طول الفئة

= ١٢ + $\frac{١٠ - ٨}{٦}$ × ٤

= ١٢ + ٠,٣٣ (٤)

= ١٢ + ١,٣٣ = ١٣,٣ كيلو جرام

$$\text{قيمة } ر_٣ = ح_د + \frac{\text{موقع } ر_٣ - \text{ك سابق}}{\text{ك الفئة الأصلية } ر_٣} \times \text{طول الفئة}$$

$$\text{موقع } ر_٣ = \frac{٤٠}{٤} \times ٣ = ٣٠$$

$$\text{قيمة } ر_٣ = ٢٤ + \frac{٣٠ - ٢٨}{٧} \times ٤$$

$$= ٢٤ + ٠,٢٩ \ (٤)$$

$$= ٢٤ + ١,١٤ = ٢٥,١٤ \text{ كيلو جرام}$$

$$\text{الانحراف الربيعى} = \frac{٢٥,١٤ - ١٣,٣}{٢} = ٥,٩٢ \text{ كيلو جرام}$$

٣- الانحراف المتوسط :

يســتخدم مقياس الانحراف المتوسط لقياس التشتت لأنه يتفادى أوجه القصور فـى المقايــيس الســابقة، حيث يستخدم الانحراف المتوسط جميع القيم أى أنه يهتم بالانحرافات لكل قيمة عن وسطها الحسابى.

مــثال : فيما يلى الدخول الشهرية (بمئات الجنيهات) لعدد من المشتغلين فى إحدى الشركات الاستثمارية والمطلوب حساب الانحراف المتوسط.

١٣	٩	١٥	١٧	٧	٨
١٤	٢٠	١٨	١٦	٥	٢

خطوات الحل :

١- احسب قيمة الوسط الحسابى للقيم جميعها.

٢- اطــرح كــل قــيمة مــن الوسط الحسابى، كقيم مطلقة بمعنى إهمال الإشارة الجبرية لها (+ ، -).

٣- اقسم اجمالى ناتج الطرح على عدد الحالات، تحصل على الانحراف المتوسط الدال على التباين فى توزيع البيانات.

حل المثال :

الدخل (س)	**$\overline{س}$ – س**
١٣	١
٩	٣-
١٥	٣
١٧	٥
٧	٥-
٨	٤-
١٤	٢
٢٠	٨
١٨	٦
١٦	٤
٥	٧-
٢	١٠-
مجـــ ١٤٤	(٥٨) مع إهمال الإشارات الجبرية

$$\overline{س} = \frac{مجـ\ س}{ن}$$

$$= \frac{١٤٤}{١٢} = ١٢$$

$$\text{الانحراف المتوسط} = \frac{مجـ |\overline{س} - س|}{ن}$$

$$= \frac{٥٨}{١٢} = ٧ \text{ جنيهات تقريبًا}$$

٤- التباين والانحراف المعيارى :

يعرف التباين بأنه متوسط مربع انحرافات القيم عن وسطها الحسابى.

وسوف نرمز للتباين بالرمز (ع٢) بينما رمزه اللاتينى (σ^2) (سيجما) كمعلم Paramater للمجتمع الأصلى والانحراف المعيارى (ع).

١- حساب التباين للبيانات غير المبوبة:

سوف يتم حساب التباين باستخدام طريقتين : الأولى باستخدام المتوسط الحسابى للقيم، والثانية هى الطريقة المباشرة.

أ - حساب التباين باستخدام المتوسط الحسابى للقيم فى المثال السابق.

س	س̄ - س	(س̄ - س)٢
١٣	١	١
٩	٣-	٩
١٥	٣	٩
١٧	٥	٢٥
٧	٥-	٢٥
٨	٤-	١٦
١٤	٢	٤
٢٠	٨	٦٤
١٨	٦	٣٦
١٦	٤	١٦
٥	٧-	٤٩
٢	١٠-	١٠٠
مجـ ١٤٤	صفر	٣٥٤

$$\text{التباين} = \frac{\text{مجـ } (\bar{\text{س}} - \text{س})^{٢}}{\text{ن}}$$

$$= \frac{٣٥٤}{١٢} = ٢٩,٥$$

الانحراف المعيارى (ع) عبارة عن الجذر التربيعى للتباين.

$$\therefore \quad ع = \sqrt{التباين}$$

$$ع = \sqrt{\frac{مجـ (\overline{س} - س)^{٢}}{ن}}$$

$$ع = \sqrt{٢٩,٥} = ٥,٤٣$$

ب- حســاب التبايــن والانحــراف المعيارى للقيم غير المبوبة باستخدام الطريقة المباشرة

س	$س^{٢}$
١٣	١٦٩
٩	٨١
١٥	٢٢٥
١٧	٢٨٩
٧	٤٩
٨	٦٤
١٤	١٩٦
٢٠	٤٠٠
١٨	٣٢٤
١٦	٢٥٦
٥	٢٥
٢	٤
١٤٤	٢٠٨٢

$$\text{التباين} = \frac{1}{\text{ن}}\left[\text{مجـ س}^{2} - \frac{(\text{مجـ س})^{2}}{\text{ن}}\right]$$

$$\text{ذ} = \frac{1}{12}\left[2082 - \frac{(144)^{2}}{12}\right]$$

$$= \frac{1}{12}\left[2082 - \frac{20736}{12}\right]$$

$$= \frac{1}{12}(2082 - 1728)$$

$$\therefore \text{ع}^{2} = \frac{354}{12} = 29{,}5$$

$$\text{ع} = \sqrt{29{,}5} = 5{,}43$$

حساب التباين والانحراف المعيارى للبيانات المبوبة:

وسوف نستخدم المعادلة التالية لحساب التباين والانحراف المعيارى.

$$\text{التباين} = \frac{\text{مجـ س}^{2}\text{ك}}{\text{مجـ ك}} - \bar{\text{س}}^{2}$$

$$\text{الانحراف المعيارى} = \sqrt{\frac{\text{مجـ س}^{2}\text{ك}}{\text{مجـ ك}} - \bar{\text{س}}^{2}}$$

وباستخدام بيانات الجدول التكرارى السابق والخاص بالأوزان، احسب قيمتى التباين والانحراف المعيارى؟

خطوات الحل :

ف	ك	س	س ك	س² ك
٤-	٣	٦	١٨	١٠٨
٨-	٥	١٠	٥٠	٥٠٠
١٢-	٦	١٤	٨٤	١١٧٦
١٦-	٤	١٨	٧٢	١٢٩٦
٢٠-	١٠	٢٢	٢٢٠	٤٨٤٠
٢٤-	٧	٢٦	١٨٢	٤٧٣٢
٢٨-٣٢	٥	٣٠	١٥٠	٤٥٠٠
مجـــ	٤٠		٧٧٦	١٧١٥٢

$\overline{س} = \frac{٧٧٦}{٤٠} =$ ١٩,٤ تقريبًا

$ع^2 = \frac{١٧١٥٢}{٤٠} - (١٩,٤)^2$

$= ٤٢٨,٨ - ٣٧٦,٣٦ = ٥٢,٤٤$

$\text{الانحراف المعيارى} = \sqrt{\frac{١٧١٥٢}{٤٠} - (١٩,٤)^2}$

$= \sqrt{٤٢٨,٨ - ٣٧٦,٣٦}$

$= \sqrt{٥٢,٤٤} = ٧,٢٤$

ويمكن للباحث طرح وسط فرضى من مراكز الفئات كتبسيط العمليات الحسابية.

مزايا الانحراف المعيارى

١- يعتبر من أدق مقاييس التشتت.

٢- يعتبر أداة تحليلية قوية فى وصف خصائص التوزيع للمجتمع الأصلى.

ولمـــا كـــان الانحـــراف المعيارى يتوقف على الوحدات المستخدمة فى قياس المشـــاهدات موضـــوع الدراسة فإن قيمته المطلقة غير مناسبة لأغراض المقارنة. ومـــن ثـــم يمكـــن الاعتماد على معامل الاختلاف كمقياس للتشتت محرر من أثر الوحدات المستخدمة فى القياس، ويمكن حساب معامل الاختلاف باستخدام المعادلة التالية.

$$\text{معامل الاختلاف} = \frac{\text{الانحراف المعيارى}}{\text{الوسط الحسابى}} \times ١٠٠$$

ثانياً: مقاييس التشتت للمتغير المتقطع :

١ – نسبة التباين The Variation Ratio:

يماثل هذا المقياس فى بساطته وسهولة حسابه مقياس المدى. ويستخدم مقياس نسبة التباين فى حالة البيانات المبوبة The Grouped data ويلائم حالات المقاييس الاسمية Nominal Scales.

تقـــيس نســـبة التباين درجة تركز الحالات حول الفئة المنوالية The Modal Category بدلا من قياس التوزيع للقيم فى جميع الفئات.

وتحسب نسبة التباين من المعادلة الآتية :

$$\text{نسبة التباين } (\text{ت}_\text{ن}) = ١ - \frac{\text{عدد الحالات فى الفئة المنوالية}}{\text{المجموع الكلى لعدد الحالات}}$$

مـــثال : باســـتخدام نســـبة التباين، وضح مدى تركز أو تشتت الحالات حول الفئة المنوالـــية فـــى توزيع ما لأحد المتغيرات. إذا كان عدد الحالات حول الفئة المنوالية ثلاث حالات، والمجموع الكلى للحالات عشرة حالات.

الحل : نسبة التباين (ت ن) $= ١ - \frac{٣}{١٠} = ٠,٧$

٢ – دليل التباين الكيفى:

يســـتخدم دليل التباين الكيفى Index of Qualitative variation للمقارنة بيـــن التبايـــن المشـــاهد للمتغـــير الاسمى، والتباين المتوقع. يتم حساب التباين

المشاهد بحساب الاختلافات فى التوزيع. بينما يمثل التباين المتوقع أقصى تشتت يمكن أن يحدث لتوزيع معين. ثم يتم حساب دليل التشتت الكيفى من المعادلة الآتية:

$$\text{دليل التباين الكيفى} = \frac{\text{التباين المشاهد}}{\text{التباين المتوقع}} \times ١٠٠$$

مثال : يشتمل الجدول الآتى على عدد المشتركين فى ندوتين علميتين من تخصصات علمية هى الخدمة الاجتماعية، علم الاجتماع، علم النفس، والانثروبولوجيا. بالإضافة إلى تباينات المشاهدة المتوقعة فى مشاركة هؤلاء الباحثين فى كل ندوة على حدة. والمطلوب قياس التباين باستخدام دليل التباين الكيفى.

التخصص للمشاركين	الندوة العلمية الأولى		الندوة العلمية الثانية	
	المشاهد	متوقع	المشاهد	متوقع
خدمة اجتماعية	٢	٥	٨	٧
علم اجتماع	١٧	٥	٦	٧
علم النفس	١	٥	٩	٧
الأنثروبولوجيا	صفر	٥	٥	٧
	ن = ٢٠		ن = ٢٨	

الحل: نلاحظ فى هذا المثال أن المتغير هنا متقطع وليس متصل.

لحساب دليل التباين الكيفى من المعادلة

$$\text{دليل التباين} = \frac{\text{التباين المشاهد}}{\text{التباين المتوقع}} \times ١٠٠$$

١- كيف يتم حساب التباين المشاهد من الجدول:

نقوم بجمع حاصل ضرب المشاهدات لكل الأزواج الممكنة منها فى كل ندوة علمية على حدة.

أ - بالنسبة للندوة الأولى:

التباين المشاهد = (٢ × ١٧) + (٢ × ١) + (٢ × صفر) + (١٧ × ١) + (١٧ × صفر) + ١ × صفر) = ٣٤ + ٢ + ١٧ = ٥٣

٢- يتم حساب التباين الأقصى بجمع كل بيانات التوزيع المشاهد (٢٠) ثم قسمتها على عدد التصنيفات التخصصية للمشاركين فى الندوة (٤) ثم يوزع الناتج من القسمة بالتساوى على كل مصنف. فالناتج من القسمة = $\frac{٢٠}{٤}$ = ٥ يصبح هو الرقم الدال على التباين المتوقع أمام كل قيمة مناظرة من التباين المشاهد.

٣- حساب التباين المتوقع مثلما تم حساب التباين المشاهد. بأنه يساوى حاصل جمع كل زوجين احتمالين من التباين المتوقع.

التباين المتوقع = [(٥ × ٥) + (٥ × ٥) + (٥ × ٥)]

+ [(٥ × ٥) + (٥ × ٥)] + ٥ × ٥ = ١٥٠

∴ نسبة التباين المشاهد إلى المتوقع يمثل دليل التباين الكيفى

= $\frac{٥٣}{١٥٠}$ × ١٠٠ = ٣٥,٣٣ (١٠٠)

= ٣٥,٣٣%

بالمثل يمكن تكرار الخطوات السابقة فى حساب دليل التباين الكيفى فى حالة الندوة العلمية الثانية وسوف نحصل على قيمة هذا الدليل وتساوى (٩٨,٣٠٥) نخلص من قيم دليل التباين الكيفى إلى أن النتائج فى هذا البحث مرضية نظرًا لأن الندوة الأولى كان التشتت محدودًا جدًا فى توزيعها ومن ثم أعطيت قيمة مئوية منخفضة للدليل بينما كان التشتت مرتفعًا فى الندوة الثانية، لذلك كانت القيمة المئوية للدليل عالية.

من فوائد دليل التباين أنه مقياس مفيد للتباين فى توزيعات البيانات المتقطعة لأنه يقيم التشتت المشاهد داخل أى توزيع فى مقابل التوزيع المتوقع.

مقاييس النزعة المركزية مقاييس والتشتت
ومستويات القياس الملائمة

		مستويات القياس			
		الأسمى	الرتبى	الفاصلة	النسبة
مقاييس النزعة المركزية	المنوال	×	×	×	×
	الوسيط		×	×	×
	المتوسط الحسابى			×	×
مقاييس التشتت	المدى		×	×	×
	التباين			×	×
	الانحراف المعيارى			×	×

المفاهيم الأساسية Key Concepts

١- التبايـــن (التشـــتت) Dispersion : يعرف بكمية أو مقدار الاختلاف أو عدم التجانس فى أى توزيع للبيانات.
٢- دلـــيل التبايـــن الكـــيفى (Index of qualitiative Variation (IQV) يعرف بمقياس التشتت للمتغيرات المتقطعة التى يتم تنظيمها فى توزيعات تكرارية.
٣- الانحراف الربيعى يعرف بنصف المسافة بين الربيع الثالث والربيع الأول.
٤- الانحراف المعيارى للعينة (ع) Sample Standard Deviation يعرف بقيمة الجذر التربيعى للتباين.
٥- تباين العينة (ع) Sample Variance يعرف بحاصل جمع كل انحرافات القيم عن الوسط الحسابى، وتربيعاتها.

تمارين

١- أكمل ما يأتى بعبارات مناسبة صحيحة:

(أ) أن نسبة مقدار التباين الذى يتم ملاحظته فعليا فى أى توزيع للقيم إلى مقدار التباين الأقصى الممكن وجوده فى هذا التوزيع يعرف ....

(ب) يعرف المدى بأنه .... بين أعلى وأقل القيم فى التوزيع.

(ج) يتجنب الانحراف الربيعى مشكلات القيم الشاذة باعتماده على .... فقط.

(د) معادلة الانحراف المتوسط هى: ....

(هـ) يرتبط الانحراف المعيارى بالتباين حيث يمثل الأول .... للثانى.

(و) كلما كان التوزيع أكثر تشتتًا، فإن قيمة التباين المعيارى تكون ....

(ز) لو كان التوزيع غير .... بالتشتت، يكون الانحراف المعيارى له ....

٢- فيما يلى عدد من الاختيارات تحت كل عبارة، والمطلوب وضع علامة (√) أمام الاختيار المناسب لهذه العبارة.

(أ) يستخدم الانحراف المعيارى فى حالة تباين أو تشتت القيم عن:

١- الوسط الحسابى.

٢- الوسيط.

٣- المنوال.

٤- التباين.

٥- جميع الاختيارات السابقة.

(ب) ما المجموعة التى تعتبر أكثر تشتتا فى قيمها من المجموعات التالية:

١-	١٠	١١	١٢	١٥	١٨	٨٤	٩٦
٢-	١٠	٤٢	٥٢	٤١	٤٦	٤٧	٩٦
٣-	١٠	٢٢	٣٩	٤٠	٥٦	٨٥	٩٦
٤-	١٠	١٠	١٠	٢٠	٢٠	٢٠	٩٦

(ج) ما المقياس الذى لا يرتبط بمقياس آخر من مقاييس التشتت الآتية :

١- الانحراف المعيارى.

٢- المدى.

٣- التباين.

٤- المتوسط الحسابى.

(د) ما الكلمة التى تستخدم فى تعريف انتشار القيم حول مقياس للنزعة المركزية :
١- المدى.
٢- التشتت.
٣- التوزيع.
٤- التباين.
٥- الاختيارات الأربعة السابقة.

٣- فــى دراسة اجتماعية أجريت على الحراك الوظيفى (الترقى) للرجال والنساء داخــل أحد الأقسام فى احدى المؤسسات الحكومية. وتوضح بيانات الجدول الآتى أن النساء تنتظر سنوات أطول فى درجاتها الوظيفية عن نظائرهن من الرجال حتى يحصلن على ترقية لدرجة أعلى وأن التفرقة بسبب تباين النوع Gender.

المطلــوب حســاب المتوسط والانحراف المعيارى للرجال والنساء. ثم اذكر رأيك حول مشكلة التفرقة على أساس النوع فى الترقى الوظيفى من خلال ما تحصل عليه من إجابات.

عدد السنوات التى يتم قضاؤها فى الأداء الوظيفى قبل الترقية الأولى	عدد العاملين	
	ذكور	إناث
١	١٣	٤
٢	٢٥	١٨
٣	٢٠	١٤
٤	١٢	١٠
٥	٣	١٠

٤- فــى دراسة أجريت على أحد السجون فى مصر للتعرف على عدد الذين يتم العفو عنهم لحسن السير والسلوك ولأسباب أخرى قبل قضائهم مدة العقوبة القانونــية. والعلاقة بين عدد هؤلاء وعدد مرات إلقاء القبض على المتهمين مــن جانــب الشــرطة. أعطــت الدراسة البيانات الموضحة بالجدول الآتى والمطلوب حساب المتوسط لهذه العينة.

عدد مرات إلقاء القبض على المتهمين	عدد الذين تم العفو عنهم من المسجونين
صفر	١٨
١	٢٢
٢	١٤
٣	٢٠
٤	١٣
٥	٧
٦	٦
	١٠٠

٥- احسب قيمة التباين للقيم الآتية :

١٠ ، ١٥ ، ١٨ ، ٢١ ، ٢٤ ، ٢٩ ، ٣٢ ، ٣٣ .



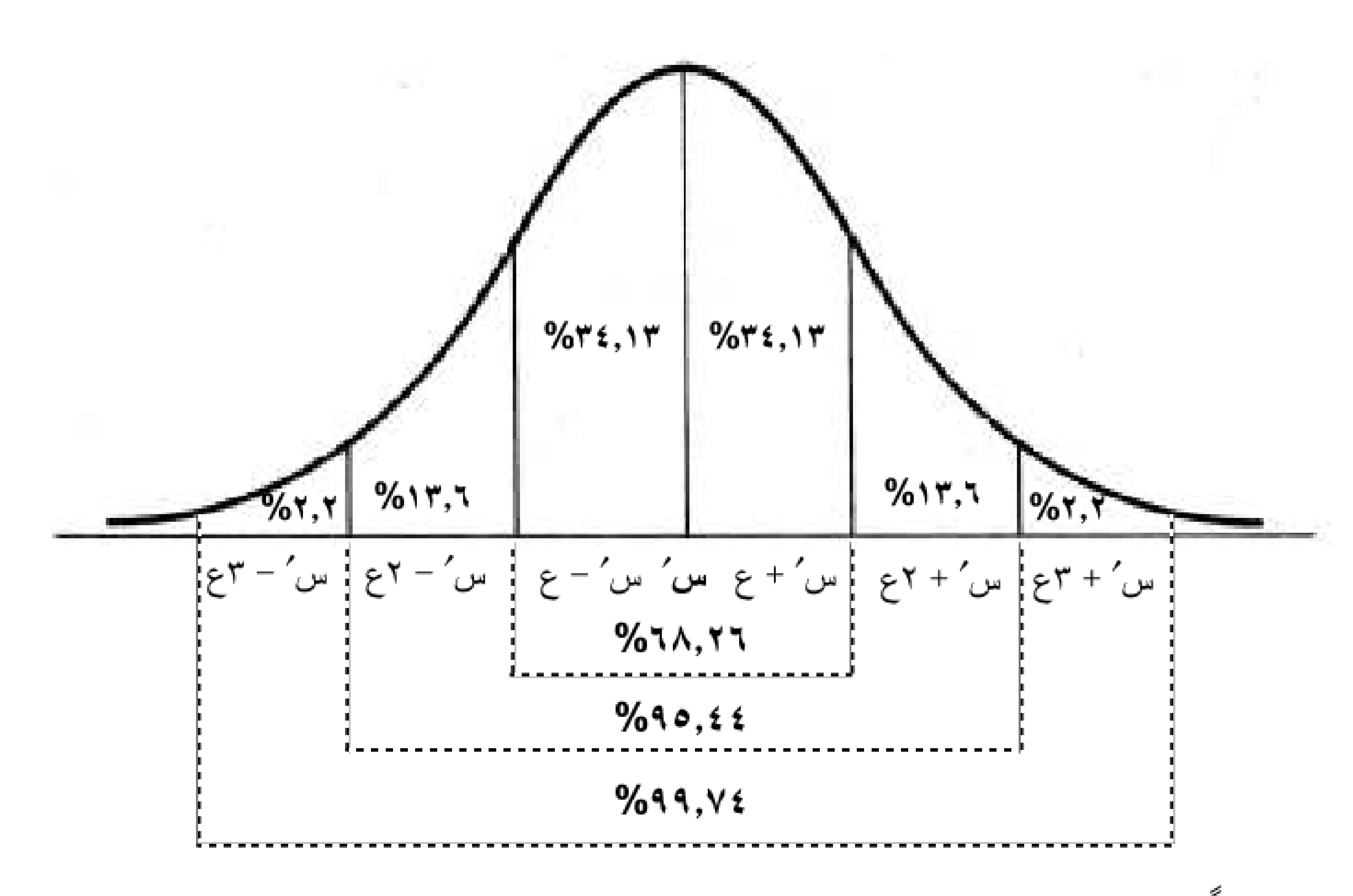
الفصل السادس
المنحنى الاعتدالى والمعايير والالتواء

أولاً: المنحنى الاعتدالى.
ثانيًا: المعايير.
ثالثًا: الالتواء.

الفصل السادس
المنحنى الاعتدالى والمعايير والالتواء

مقدمة :

يعتبر المنحنى الاعتدالى من أهم التوزيعات الاحتمالية لأن معظم الظواهر فى حياتنا تتبع ذلك التوزيع مثل الأوزان والأطوال ومقاييس الذكاء لمجموعة كبيرة من الأفراد.



أولاً: خواص المنحنى الاعتدالى: (اعتماد علام ويسرى رسلان، ١٩٩٢: ١٦٠-١٦١).

١- إن وسطه الحسابى = صفر والانحراف = ١

٢- له قمة منوالية واحدة وطرفاه يمتدان إلى ما لا نهاية حيث يقتربان من المحور الأفقى ولكن لا يلتقيان به أبدًا، أى مفتوح عند الطرفين.

٣- يشبه جرس المدرسة Bell-Shaped.

٤- له محور تماثل يمر بالقمة ويقطع المحور الأفقى عند النقطة التى تحدد الوسط الحسابى ومحور التماثل هذا بقسم المنحنى إلى قسمين متساويين فى المساحة.

٥- إنـــه توزيع معيارى Standard منتظم Symmetric بمعنى أن المساحة أسفل المنحـــنى تنقســم إلى ستة أقسام ثابتة من حيث المساحة مهما اختلف مستوى

القـياس أو معياره حيث ينقسم إلى ثلاثة أقسام على يمين المتوسط الحسابى، وثلاثة على يسار المتوسط الحسابى.

٦- إنـه يعكس العلاقة الرياضية بين الوسط الحسابى والانحراف المعيارى حيث تحـدد هذه العلاقة التوزيع الاحتمالى للقيم فمثلاً ٣٤,١٣% من المساحة أسفل المنحـنى تقع بين صفر + ١ع ومثلها بين صفر - ١ع وهذا يعنى أن احتمال وقـوع قـيمة ما من قيم التوزيع بين صفر وواحد هو ٣٤,١٣% على جانبى الوسـط الحسـابى ويتحدد موقعها يمينًا أو يسارًا بمقارنتها بالوسط الحسابى للـتوزيع.. بمعنى إذا كانت أكبر من الوسط الحسابى تكون على اليمين أما إذا كانت أقل منه فهى على اليسار.

ثانيًا : المعايير :

إن أى درجـة خام Raw score ليس لها أى دلالة ولا تستعمل فى المقارنات لأن هـذه الدرجة الخام ليس لها أى معنى. فإذا فرضنا أن أحد الطلاب حصل على سـبعين درجـة ١٠٠/٧٠ فـى مادة مبادئ الاحصاء فلا نستطيع الحكم على هذا الطالب أنه قوىٌ فى تحصيله لهذه المادة أم ضعيفُ فقد يكون الاختبار صعبًا فتكون هـذه الدرجـة أعلى الدرجات، أو قد يكون الامتحان سهلاً فتكون هذه الدرجة أقل الدرجـات أو تكون هذه الدرجة متوسطة، ومن ثم لابد أن تنسب هذه الدرجة الخام إلـى المتوسـط الحسـابى لدرجـات المجموعـة التى تنتمى إليها درجة الطالب والانحـراف المعـيارى لهـذه الدرجات. ومن ثم يتم تحويلها إلى درجة معيارية standard score وفـى ضـوء هـذه الدرجة المعيارية يمكن الحكم على مستوى الطالب إذا كان مساوى للمتوسط الحسابى أو أعلى أو أقل منه.

١- الدرجة المعيارية Standad Score :

ويمكن حساب الدرجة المعيارية لأى درجة خام على أساس حساب الفرق بين هـذه الدرجـة والمتوسـط الحسـابى مقسومًا على الانحراف المعيارى لدرجات المجموعة.

$$\text{الدرجة المعيارية} = \frac{\text{القيمة} - \text{المتوسط الحسابى}}{\text{الانحراف المعيارى}}$$

وباستخدام الرموز تكون المعادلة على النحو التالى:

الدرجة المعيارية ويرمز لها بالرمز Z .

القيمة يرمز لها بالرمز س.

المتوسط الحسابى يرمز له بالرمز س َ.

الانحراف المعيارى يرمز له بالرمز ع.

$$Z = \frac{\text{س} - \text{س}'}{\text{ع}}$$

مثال :

احسب الدرجة المعيارية لثلاث طالبات فى مادة مبادئ الإحصاء علمًا بأن المتوسط الحسابى لتوزيع الدرجات ٥٠ درجة من المجموع الكلى للدرجات (١٠٠ درجة) والانحراف المعيارى يساوى ١٠ درجات. المطلوب تحديد الدرجة المعيارية لكل طالبة على حدة وموقعها على المنحنى الاعتدالى.

درجات	الطالبات
أسماء	٦٠ درجة
منال	٤٠ درجة
نادين	٥٠ درجة

الحــل :

لابد من إيجاد الدرجات المعيارية لهذه الدرجات الخام أولاً ثم يتم تحديد موقع كل درجة فى مجال الانحراف المعيارى أسفل المنحنى.

$$Z_{\text{أسماء}} = \frac{٦٠ - ٥٠}{١٠} = ١$$

$$Z_{\text{منال}} = \frac{٤٠ - ٥٠}{١٠} = -١$$

$$Z_{\text{نادين}} = \frac{٥٠ - ٥٠}{١٠} = \text{صفر}$$

أمـــا بالنســـبة لأســـماء فقد حصلت على درجة (٦٠)، أى أعلى من المتوسط الحســـابى بمقـــدار ١ع ومـــن ثـــم تقـــع فـــى مجـــال الانحراف المعيارى الأول (٣٤,١٣%) الموجـــب علـــى يميـــن الوسط الحسابى أما درجة منال فهى أقل من المتوســـط الحســـابى بمقـــدار واحد انحراف معيارى ومن ثم فهى تقع ض مجال الانحراف المعيارى الأول السالب على يسار المتوسط الحسابى، أما بالنسبة لدرجة نادين فهى مساوية لمتوسط المجموعة.

∴ الدرجة الخام = المتوسط ± الدرجة المعيارية × ع

٢- الدرجة التائية:

هـــى عـــبارة عن درجة معيارية متوسطها ٥٠ وانحرافها المعيارى ١٠ وبها يمكـــن الـــتخلص من الإشارات السالبة والموجبة فى الدرجة المعيارية (محمود أبو النيل، ١٩٩٨)(*).

فإذا كان لدينا درجة معيارية - ٢

فإن الدرجة التائية المقابلة لها تساوى

المتوسط ± الدرجة المعيارية × الانحراف المعيارى

∴ الدرجة التائية = ٥٠ - (٢ × ١٠)

= ٥٠ - ٢٠ = ٣٠

ثالثاً : الالتواء :

كمـــا سبق أن أوضحنا فى الفصل الأول فإن الالتواء عبارة عن خاصية تعنى ابـــتعاد الـــتوزيع التكرارى عن التماثل ويمكن قياس الالتواء Skewness باستخدام معادلة كارل بيرسون من خلال العلاقة الآتية :

المتوسط الحسابى - المنوال = ٣ (المتوسط الحسابى - الوسيط)

---

* ويمكن معرفة هل هناك فرق له دلالة إحصائية بين درجـــة الطالبة الخـــام ومتوسط المجموعة باستخدام الدرجـــة المعـــيارية، ويعتـــبر الفرق دالاً عند مستوى ٠,٠٥ إذا كانت الدرجة المعيارية ١,٩٦ ودالاً عند مستوى ٠,٠١ عندما تساوى ٢,٥٨ انظر: محمود السيد أبو النيل الإحصاء النفسى والاجتماعى والتربوى، المؤسسة الابراهيمية، ١٩٩٨، ص ١٣٨.

وقد اقترح بيرسون مقياسين لحساب الالتواء هما:

أ – معامل بيرسون الأول للالتواء = $\frac{\text{المتوسط الحسابى – المنوال}}{\text{الانحراف المعيارى}}$

ب– معامل بيرسون الثانى للالتواء = $\frac{\text{٣ (المتوسط الحسابى – الوسيط)}}{\text{الانحراف المعيارى}}$

ويجـــدر التنويه أن القسمة على الانحراف المعيارى تهدف إلى تحويل معامل الالتواء إلى مقياس نسبى يمكن به مقارنة الالتواء فى التوزيعات المختلفة.

مثال:

كان المتوسط الحسابى والمنوال والانحراف المعيارى لمجموعة من القيم كما يلى:

المتوســـــط = ٣٣,١٧٥

المنـــــــوال = ٣٢,٠٦

الانحراف الميعارى = ٥,٥٦

والمطلوب حساب الالتواء باستخدام معامل بيرسون الأول للالتواء:

معامل الالتواء = $\frac{\text{المتوسط الحسابى – المنوال}}{\text{الانحراف المعيارى}}$

= $\frac{\text{٣٢,٠٦ – ٣٣,١٧٥}}{\text{٥,٥٦}}$ = ٠,٢

ويلاحــظ أن الالـــتواء موجب الإشارة أى متجه نحو اليمين ويمكن باستخدام خاصـــية الالتواء تحديد مدى تماثل التوزيع التكرارى. فالتوزيعات التكرارية تكون متماثلة عندما يساوى معامل الالتواء صفرًا.



الفصل السابع

الارتبـاط

مقدمة

الارتباط البسيط ومعاملاته.

١- معامل بيرسون.

٢- معامل سبيرمان.

٣- معامل فاى.

٤- معامل التوافق.

٥- الارتباط الجزئى والمتعدد.

الفصل السابع
الارتباط

مقدمة :

يستخدم معامل الارتباط في البحوث الوصفية التي تهدف إلى وصف درجة العلاقة بين المتغيرات وصفاً كميًا ويعبر عن درجة العلاقة بين المتغيرات بمعامل الارتباط بمعنى أن درجات متغير ما ترتبط بدرجات متغير آخر. ويتراوح حجم معامل الارتباط بين (-١ ، +١) . فكلما اقترب مقدار المعامل من (١) تكون العلاقة قوية بين المتغيرين ، وكلما انخفض المعامل عن (١) فإن العلاقة تضعف تدريجيًا. ومن ثم فإن معامل الارتباط يحدد حجم العلاقة بين المتغيرين واتجاه هذه العلاقة.

فإذا كانت العلاقة موجبة (+) أي طردية Positive Correlation فإنها تعني أنه إذا زاد المتغير (س) قابلته زيادة في المتغير (ص) وإذا انخفضت الدرجة على المتغير (س) قابلها انخفاض على المتغير (ص) . أما العلاقة السالبة (العكسية) Negative Correlation فتعني أن الزيادة في أحد المتغيرين يقابلها انخفاض في المتغير الثاني والعكس بالعكس.

مثال :

إذا أراد باحث أن يكشف عن حجم العلاقة واتجاهها بين الدخل الشهري (س) لعينة من الأسر والاستهلاك الشهري (ص) لهذه الأسر فإنه يستخدم معامل الارتباط لتحديد هذه العلاقة واتجاهها.

أنواع الارتباط:

ينقسم الارتباط إلى الأنواع التالية :

١- الارتباط الخطي البسيط Simple Linear Correlation ويقيس معامل الارتباط العلاقة بين متغيرين أحدهما تابع (ص) والأخر مستقل (س).

أ - الارتباط البسيط ومعاملاته هي :

١- معامل بيرسون.

٢- معامل سبيرمان.

٣- معامل فاي.

٤- معامل التوافق.

وفى إطار ما سبق يهدف الفصل إلى أن:

١- يعــرف الدارس معنى الارتباط وأن يحدد بمجرد النظر إلى الشكل الانتشارى الذى يوضح العلاقة بين المتغيرين اتجاه العلاقة ومدى قوتها.

٢- يكشــف عن حجم العلاقة بين المتغيرين (س) المتغير المستقل، (ص) المتغير التابع. كما يستطيع أن يحدد اتجاه العلاقة (طردية أم عكسية).

٣- يعــرف كيفــية حساب معاملات الارتباط المختلفة (بيرسون، سبيرمان، فاى) باستخدام المعادلة الرياضية لكل معامل منها.

أولاً : حساب معامل بيرسون للارتباط (ر) من القيم الخام:

يستخدم معامل بيرسون للارتباط في التعرف على العلاقة بين المتغيرات ذات البيانات الفاصلة والنسبية Interval and Ratio.

١- معامل بيرسون للارتباط :

وتقــوم طــريقة بيرســون فــي حســاب معــامل الارتــباط بين متغيرين (س ، ص) علــى اســتخدام انحــرافات القيم عن المتوسط الحسابي لكل منهما أي الفــرق بين $(\overline{\text{س}} - \text{س})$ للمتغير الأول ، $(\overline{\text{ص}} - \text{ص})$ للمتغير الثاني . وذلك على أســاس أن الارتــباط يقيس العلاقة بين التغير في قيم (س) والتغير في قيم (ص). وتعتبر طريقة قياس انحرافات القيم عن متوسطها الحسابي هي أفضل طرق لقياس هــذا التغير وتحسب قيمة معامل الارتباط بين المتغيرين (س) ، (ص) من المعادلة الآتية بفرض أن ن = حجم العينة.

$$\text{ر} = \frac{\text{ن مجـ س ص} - (\text{مجـ س})(\text{مجـ ص})}{\sqrt{[\text{ن مجـ س}^{٢} - (\text{مجـ س})^{٢}]\,[\text{ن مجـ ص}^{٢} - (\text{مجـ ص})^{٢}]}} \quad (١)$$

مثال :

فــيما يلــى بيانات حول الدخل الشهرى (بمئات الجنيهات) (س) والاستهلاك (بمئات الجنيهات) (ص) لسبع أسر.

س	٨	١٠	١٢	١٢	١٣	١٥	٢٠
ص	٨	٩	١٢	١٠	١٠	١٣	١٩

المطلوب: حساب معامل بيرسون للارتباط.

الحل :

س	ص	س^{٢}	ص^{٢}	س ص
٨	٨	٦٤	٦٤	٦٤
١٠	٩	١٠٠	٨١	٩٠
١٢	١٢	١٤٤	١٤٤	١٤٤
١٢	١٠	١٤٤	١٠٠	١٢٠
١٣	١٠	١٦٩	١٠٠	١٣٠
١٥	١٣	٢٢٥	١٦٩	١٩٥
٢٠	١٩	٤٠٠	٣٦١	٣٨٠
٩٠	٨١	١٢٤٦	١٠١٩	١١٢٣

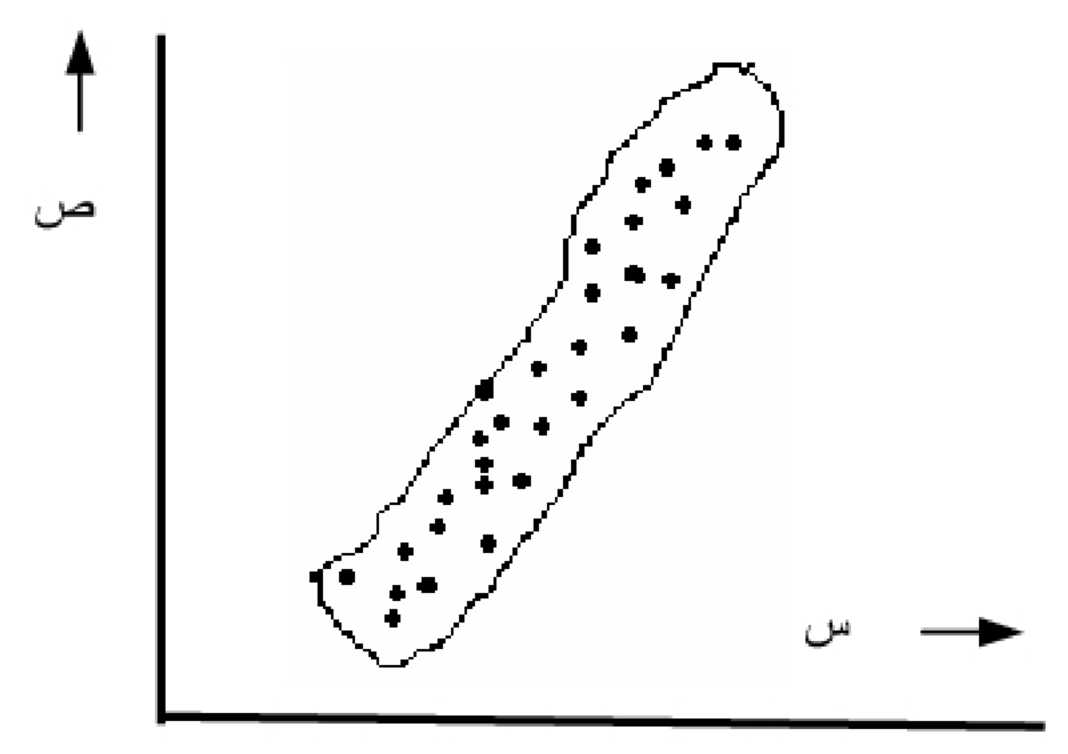
وبالتعويض باستخدام المعادلة السابقة لإيجاد قيمة (ر) فإن:

$$ر = \frac{٧ \times ١١٢٣ - (٩٠)(٨١)}{\sqrt{[٧ \times ١٢٤٦ - (٩٠)^{٢}][٧ \times ١٠١٩ - (٨١)^{٢}]}}$$

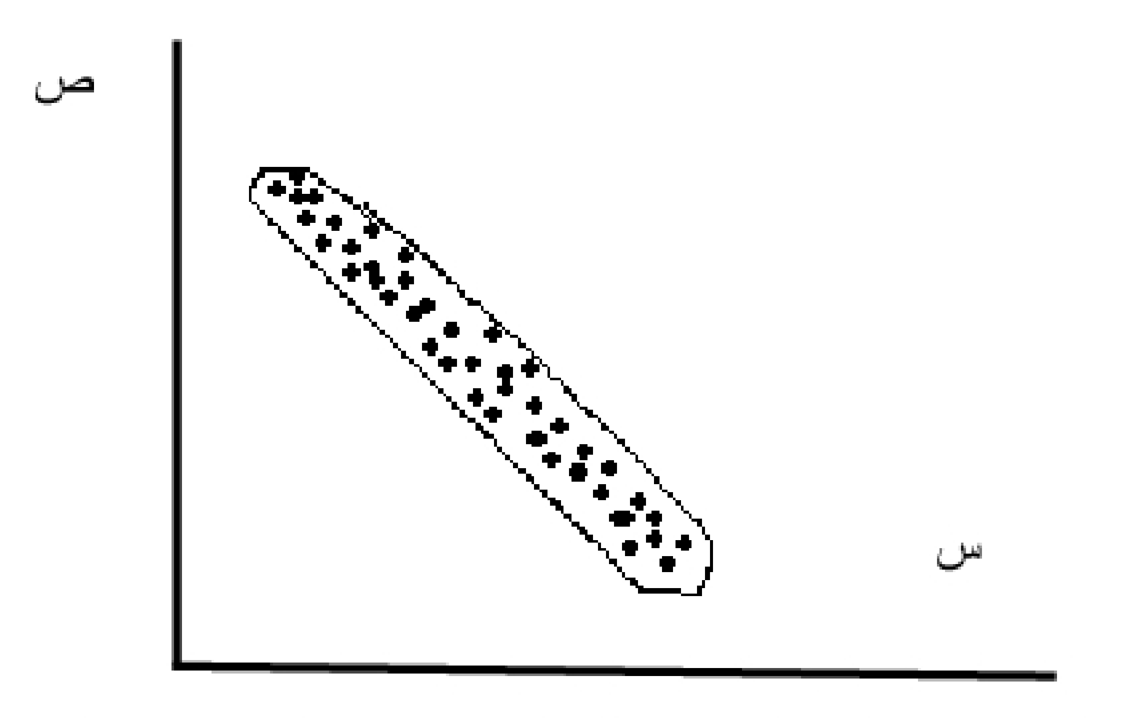
$$= \frac{٥٧١}{٥٩٦,٤٨}$$

$= ٠,٩٦$ (علاقة طردية قوية)

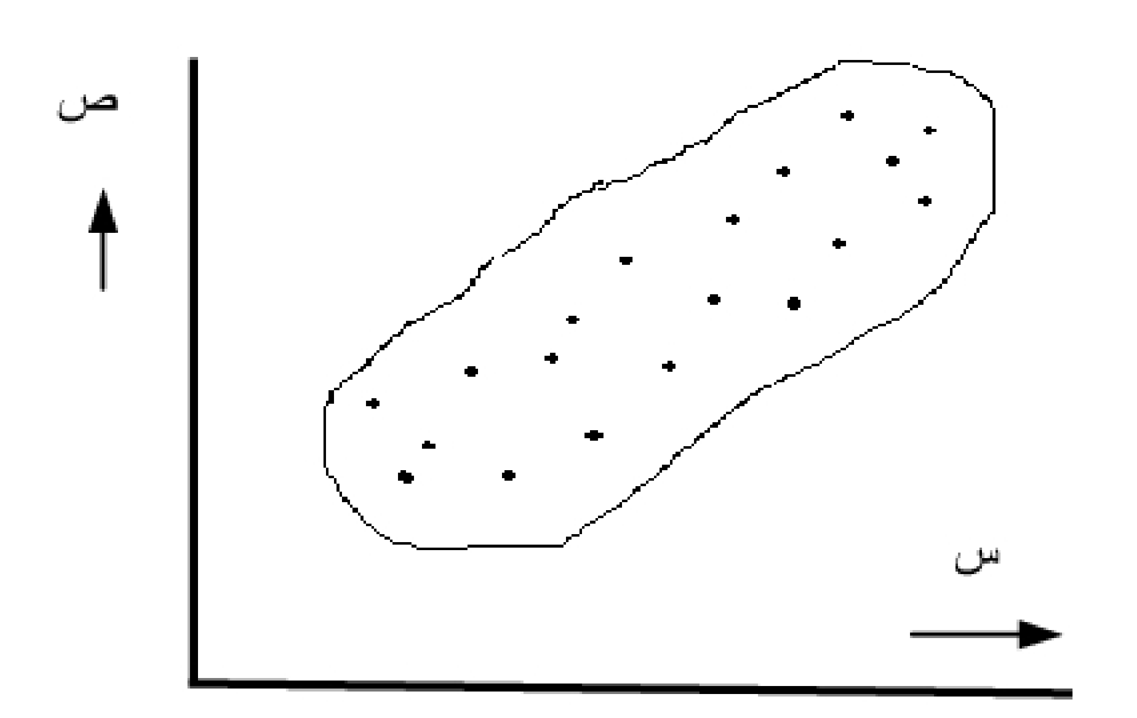
بعـــد حساب قيمة معامل الارتباط، يمكن استخدام الشكل الانتشارى فى تمثيل هـــذه المعادلة بيانياً. وكلما ازدادت مساحة شكل المنحنى قلت قوة العلاقة، بينما يتم تحديد اتجاهها ويتضح ذلك من الأشكال الآتية رقم (٧-١، ٢، ٣).



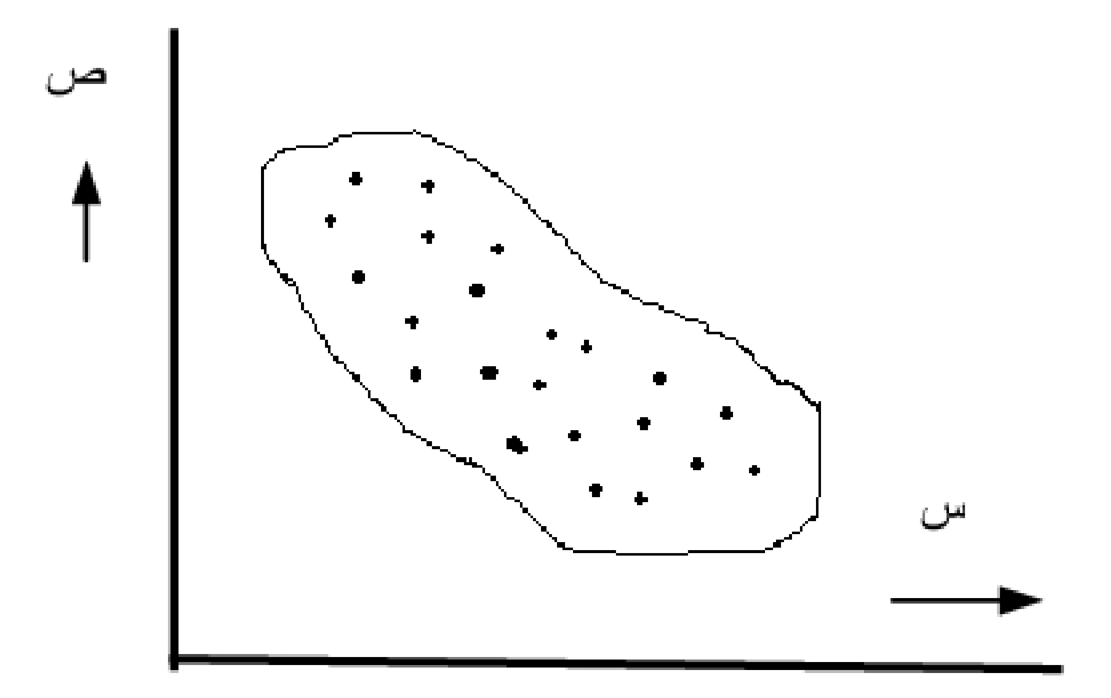
شكل رقم (٧-١) علاقة ارتباطية طردية قوية



شكل رقم (٧-٢) علاقة ارتباطية عكسية قوية



شكل رقم (٧-٣) علاقة ارتباطية طردية ضعيفة



شكل رقم (٧-٤) علاقة ارتباطية عكسية ضعيفة

جـــدول العلاقـــة بيـــن حجم معامل الارتباط (حدود تقريبية) ودرجة العلاقة الارتباطية بين المتغيرين (س، ص):

حجم (ر)	درجة العلاقة الارتباطية
(٠,٧٥ إلى ١) (-٠,٧٥ إلى - ١)	علاقـــة ارتباطـــية قويـــة جـــدًا (طردية وعكسية)
(٠,٥٠ إلى ٠,٧٤) (-٠,٥٠ إلى -٠,٧٤)	علاقـــة ارتباطـــية قويـــة (طردية وعكسية)
(٠,٢٥ إلى ٠,٤٩) (-٠,٢٥ إلى - ٠,٤٩)	علاقـــة ارتباطـــية متوســـطة (طردية وعكسية)
(صفر إلى ٠,٢٤) (صفر إلى -٠,٢٤)	علاقـــة ارتباطـــية ضـــعيفة (طردية وعكسية)

ويتسم معامل الارتباط بخاصيتين يستفاد منهما في حسابه:

الأولـــى هي أنه إذا طرحنا أو (جمعنا) رقم ثابت من جميع قيم (س) وثابت آخر من جميع قيم (ص) فإن قيمة معامل الارتباط لا تتغير .

الخاصـــية الثانية تشير إلى أن قيمة (ر) لا تتغير إذا قسمنا أو (ضربنا) جميع قيم (س) على ثابت وأيضا جميع قيم (ص) على ثابت آخر .

ولتوضـــيح ذلـــك سوف نعطي مثالا على الخاصية الأولى وهي طرح وسط فرضي من قيم (س) ثم وسط فرضي آخر من قيم (ص) .

مثال :

يوضح الجدول الآتي قيم (س) و(ص) والمطلوب حساب معامل الارتباط من القيم الأصلية ثم استخدام (٣٩) كوسط فرضي لقيم (س)، (٣٥) كوسط فرضي لقيم الظاهرة (ص).

س	٢٥	٤٢	٣٥	٣٧	١٥	٢٤	٤٣	٥٣	٤٧	٣٩
ص	٢٢	٢٧	٤٥	٣٥	٣٣	٣٠	٣٢	٤٤	٤٥	٢٧

المطلوب :

١- حساب معامل الارتباط (ر) من القيم الأصلية

قيم س	قيم ص	س^٢	ص^٢	س ص
٢٥	٢٢	٦٢٥	٤٨٤	٥٥٠
٤٢	٢٧	١٧٦٤	٧٢٩	١١٣٤
٣٥	٤٥	١٢٢٥	٢٠٢٥	١٥٧٥
٣٧	٣٥	١٣٦٩	١٢٢٥	١٢٩٥
١٥	٣٣	٢٢٥	١٠٨٩	٤٩٥
٢٤	٣٠	٥٧٦	٩٠٠	٧٢٠
٤٣	٣٢	١٨٤٩	١٠٢٤	١٣٧٦
٥٣	٤٤	٢٨٠٩	١٩٣٦	٢٣٣٢
٤٧	٤٥	٢٢٠٩	٢٠٢٥	٢١١٥
٣٩	٢٧	١٥٢١	٧٢٩	١٠٥٣
٣٦٠	٣٤٠	١٤١٧٢	١٢١٦٦	١٢٦٤٥

ر = ١٠ (١٢٦٤٥) − (٣٦٠) (٣٤٠) ÷ √([١٠ (١٤١٧٢) − (٣٦٠)^٢] [١٠ (١٢١٦٦) − (٣٤٠)^٢])

= ٠,٤٧

٢- استخدام الوسط الفرضي :

س	ص	ح س س − ٣٩	ح ص ص − ٣٥	ح^٢ س	ح^٢ ص	ح س	ح ص
٢٥	٢٢	-١٤	-١٣	١٩٦	١٦٩	١٨٢	
٤٢	٢٧	٣	-٨	٩	٦٤	-٢٤	
٣٥	٤٥	-٤	١٠	١٦	١٠٠	-٤٠	
٣٧	٣٥	-٢	صفر	٤	صفر	صفر	
١٥	٣٣	-٢٤	-٢	٥٧٦	٤	٤٨	
٢٤	٣٠	-١٥	-٥	٢٢٥	٢٥	٧٥	
٤٣	٣٢	+٤	-٣	١٦	٩	-١٢	
٥٣	٤٤	١٤	٩	١٩٦	٨١	١٢٦	
٤٧	٤٥	٨	١٠	٦٤	١٠٠	٨٠	
٣٩	٢٧	صفر	-٨	صفر	٦٤	صفر	
مجـ		-٣٠	-١٠	١٣٠٢	٦١٦	٤٣٥	

$$\bar{\text{ح}}_{\text{س}} = \frac{\text{مجـ ح}_{\text{س}}}{\text{ن}} = \frac{\text{٣٠-}}{\text{١٠}} = \text{٣-}$$

$$\bar{\text{ح}}_{\text{ص}} = \frac{\text{مجـ ح}_{\text{ص}}}{\text{ن}} = \frac{\text{٣٠-}}{\text{١٠}} = \text{١-}$$

$$\text{ع س} = \frac{\text{مجـ ح}^{\text{٢}}_{\text{س}}}{\text{ن}} - (\bar{\text{ح}}_{\text{س}})^{\text{٢}}$$

حيث ع س هي الانحراف المعياري (س)

$$= \sqrt{\frac{\text{١٣٠٢}}{\text{١٠}} - (\text{- ٣})^{\text{٢}}} = \text{١١,٠١}$$

$$\text{ع}_{\text{ص}} = \sqrt{\left(\frac{\text{مجـ ح}_{\text{ص}}}{\text{١٠}}\right) - (\bar{\text{ح}}_{\text{ص}})^{\text{٢}}}$$

حيث ع ص = الانحراف المعياري (ص)

$$= \sqrt{\frac{\text{٦١٦}}{\text{١٠}} - (\text{-١})^{\text{٢}}}$$

$$= \sqrt{\text{٦١,٦ - ١}} = \sqrt{\text{٦٠,٦٠}} = \text{٧,٧٨}$$

وباســتخدام المعادلة الآتية يحسب معامل الارتباط بطرح ثابت من قيم (س)، ثابت آخر من قيم (ص) (وسط فرضي) .

$$= \frac{\text{٤٣٥} - (\text{-٣} \times \text{-١})}{\text{١١,٠١} \times \text{٧,٧٨}}$$

$$= \frac{\text{٤٣,٥٠ -٣}}{\text{١١,٠١ - ٧,٧٨}}$$

$$= \frac{\text{٤٠,٥٠}}{\text{٥٨,٦٦}} = \text{٠,٤٧}$$

أمـــا فـــي حالة تقدير معامل بيرسون للارتباط في الجدول المزدوج ، يفضل اســتخدام حزمة البرامج الإحصائية للعلوم الاجتماعية Statistical Package for Social Science (SPSS) . لما تتسم به من سرعة ودقة عالية.

معامل ارتباط سبيرمان للرتب: Spearman Rank Order Correlation Coefficient

يستخدم هذا المعامل في حالة العينات صغيرة الحجم ويعتمد على ترتيب القيم فـــي كـــل متغير موضوع الدراسة . وفي نفس الوقت يعتبر معامل سبيرمان حالة خاصة من معامل ارتباط بيرسون لمتغيرين كلاهما يتم قياسه بالمقاييس الرتبية فلو فرضـــنا أن باحثا يريد أن يعرف العلاقة بين حجم الفصل الدراسي للفرقة النهائية في اثنى عشر كلية جامعية لعام معين وليكن عام ٢٠٠٨-٢٠٠٩ ونسبة الخريجين ممـــن يســـتكملون دراســـاتهم العلـــيا للماجستير والدكتوراه وحصل على البيانات الموضحة بالجدول التالى.

الكلية	حجم الفصل الدراسى (س)	نسبة الدارسين (ص)
أ	٣٠٦٨	٢,٩
ب	٢٥٨٤	٣,٦
ج	٢٠٦٧	١,٣
د	١٥٨٤	٦,٨
هـ	١٠٩٣	٤,٩
و	٨٤٧	١,٨
ز	٦٩٨	٤,٣
ح	٥٦٣	٨,٦
ط	٣٩٨	٥,٧
ى	٣٠٤	٨,٩
ك	٢١٨	٤,٧
ل	١٣٠	٧,٥

خطوات الحل :

١- نقـــوم بترتيب المتغير الأول (س) ترتيبًا تصاعديًا أو تنازليًا فعلى سبيل المثال فـــي حالـــة الترتيـــب التصاعدي يتم إعطاء الرتبة الأولى لأقل درجة والرتبة الثانية للدرجة التي تليها، وهكذا ويوضع ذلك في عمود رتبة الظاهرة (س).

٢- نقـــوم بترتيب المتغير الثاني (ص) بنفس طريقة ترتيب المتغير (س) ويوضع في العمود الثاني من نفس خانة ترتيب الظاهرة بالجدول.

٣- نقوم بحساب الفرق بين رتبة (س) ورتبة (ص) وذلك بطرح رتبة الثاني من رتبة الأول أو العكس. ويوضع الناتج فى العمود المسمى بالفرق بين الترتيبين ويرمز للفرق بالرمز (ف) .

٤- نقـــوم بعـــد ذلك بتربيع الفرق ($ف^٢$) ويوضع في العمود الثاني المسمى الفرق بين الترتيبين ومربعه.

٥- نقوم بجمع القيم الموجودة في العمود ($ف^٢$) لإيجاد مجـــ $ف^٢$.

٦- يتم تطبيق معامل ارتباط الرتب لسبيرمان وصيغته كالأتي :

$$\text{معامل سبيرمان للرتب} = ١ - \frac{٦ \text{ مجـــ } ف^٢}{ن (ن^٢ - ١)}$$

حل المثال :

الكلية	حجم الفصل (س)	نسبة الدارسين (ص)	ترتيب س	ترتيب ص	الفروق رتبة س، رتبة ص (ف)	مربع الفروق $ف^٢$
أ	٣٠٦٨	٢,٩	١	١٠	٩-	٨١
ب	٢٥٨٤	٣,٦	٢	٩	٧-	٤٩
جـ	٢٠٦٧	١,٣	٣	١٢	٩-	٨١
د	١٥٨٤	٦,٨	٤	٤	صفر	صفر
هـ	١٠٩٣	٤,٩	٥	٦	١-	١
و	٨٤٧	١,٨	٦	١١	٥-	٢٥
ز	٦٩٨	٤,٣	٧	٨	١-	١
ح	٥٦٣	٨,٦	٨	٢	٦	٣٦
ط	٣٩٨	٥,٧	٩	٥	٤	١٦
ي	٣٠٤	٨,٩	١٠	١	٩	٨١
ك	٢١٨	٤,٧	١١	٧	٤	١٦
ل	١٣٠	٧,٥	١٢	٣	٩	٨١
مجـــ					صفر	٤٦٨

وباستخدام المعادلة رقم (٧-٨) يكون الحل كالاتى :

$$\text{معامل سبيرمان} = ١ - \frac{٦ \times ٤٦٨}{١٢ (١٤٤-١)} = ٠,٦٣٦$$

ويلاحـــظ أن المعادلة السابقة لسبيرمان قد تم استنتاجها من معادلة قيمة معامل الارتباط (ر) بشرط افتراضية :

١- تساوى الوسطين الحسابيين للمتغيرين (س، ص).
٢- تساوى الانحرافين المعياريين للمتغيرين.
٣- ترتيب القيم الأصلية ترتيبًا تصاعديًا أو تنازليا.
٤- مـــن عـــيوب معـــامل ارتباط الرتب أنه إذا تغيرت القيم لن يتأثر قيم معامل الارتـــباط لكنه فى حالة معامل ارتباط بيرسون فإن أى تغيير في القيم سوف يؤثر على قيم معامل الارتباط.

أمـــا فـــى الحـــالات الـــتى يجد فيها الباحث قيما متشابهة لأى من المتغيرين (س،ص) فـــيقوم بإعطـــاء ترتيـــب متوسط للقيم المتشابهة فى حالة حساب معامل ارتباط الرتب كما يتضح ذلك من المثال التالى.

مثال :

أراد باحـــث اجـــتماعى أن يدرس العلاقة الارتباطية بين عدد الخريجين فى قســـم اللغـــة الفرنسية بإحدى كليات الآداب وفرص العمل المتاحة لهم فى مدينة القاهرة وذلك من خلال حصر عدد المشتغلين منهم خلال سبع سنوات متتالية بدءًا من عام ٢٠٠٠ حتى عام ٢٠٠٦ وأعطت الدراسة البيانات الموضحة فى الجدول رقم (٦-١).

السنة	عدد الخريجين سنويًا(ص)	عدد المشتغلين بالقاهرة (س)
٢٠٠٠	٧٤	١٨
٢٠٠١	٦٥	٩
٢٠٠٢	٧٩	١٢
٢٠٠٣	٨٠	١٥
٢٠٠٤	٧٩	١٢
٢٠٠٥	٨٢	١٦
٢٠٠٦	٨٦	١٧

تلاحــظ مــن البيانات السابقة تساوى رقمى السنتين ٢٠٠٢ ، ٢٠٠٤ للمتغير (ص) ومن ثم نأخذ المتوسط الحسابى للترتيب $\frac{٣ + ٤}{٢}$ = ٣,٥ فتأخذ كل منهما فى الترتيــب قيمة واحدة هى (٣,٥) وكذلك الحال بالنسبة للمتغير (س) فنأخذ المتوسط الحسابى للترتيب $\frac{٢ + ٣}{٢}$ = ٢,٥

حل المثال :

	س	ص	رتبة س	رتبة ص	ف	ف٢
٢٠٠٠	١٨	٧٤	٧	٢	٥	٢٥
٢٠٠١	٩	٦٥	١	١	صفر	صفر
٢٠٠٢	١٢	٧٩	٢,٥	٣,٥	١-	١
٢٠٠٣	١٥	٨٠	٤	٥	١-	١
٢٠٠٤	١٢	٧٩	٢,٥	٣,٥	١-	١
٢٠٠٥	١٦	٨٢	٥	٦	١-	١
٢٠٠٦	١٧	٨٦	٦	٧	١-	١
	مجــ				صفر	٣٠

معامل سبيرمان = ١- $\frac{٦ \times ٣٠}{٧ (٤٩ - ١)}$ = ٠,٤٦

معامل الارتباط فاى : φ The Phi Coefficient

وهــو حالــة خاصــة مــن معــامل بيرسون للارتباط ويستخدم عندما يكون المتغــيران ذا بــيانات اسمية وتكون البيانات لكل منهما من النوع الاسمى الثنائى النوع (ذكور، إناث)، الحزب السياسى (وطنى ، ومعارض).

مثال :

أراد باحــث اجتماعى أن يحدد العلاقة بين النوع (ذكر وأنثى) لعينة من المبحوثين وانضمامهم لحزب سياسي معين وحيث إن المتغيرين من النوع الاسمى فمــن الأجــدر أن نستخدم تمييزًا رقميًا للنوع كأن يعطى للإناث رقم (١) وللذكور (صفر) ويكــرر نفس العمل بالنسبة للحزبين السياسيين كأن يعطي للحزب الوطني

(١) والحزب المعارض (صفرًا) والجدول التالى يتضمن البيانات التي حصل عليها الباحث للمتغيرين س ، ص.

المفردات	النوع	الحزب السياسي			
	س	ص	$س^٢$	$ص^٢$	س ص
أ	١	١	١	١	١
ب	١	١	١	١	١
جـ	١	صفر	١	صفر	صفر
د	١	١	١	١	١
هـ	١	١	١	١	١
و	صفر	صفر	صفر	صفر	صفر
ز	صفر	١	صفر	١	صفر
ح	صفر	١	صفر	١	صفر
ط	صفر	صفر	صفر	صفر	صفر
ي	صفر	صفر	صفر	صفر	صفر
مجـ	٥	٦	٥	٦	٤

وبتطبيق معادلة بيرسون للارتباط يكون حل المثال على النحو الأتي:

$$ر = \frac{١٠ (٤) - (٥) (٦)}{\sqrt{[(١٠) (٥) - (٥)^٢] [(١٠) (٦) - (٦)^٢]}} = ٠,٤٠٩$$

ويـــدل هـــذا المعـــامل على وجود علاقة ارتباطية طردية متوسطة بين النوع والحـــزب السياســـي الذي ينتمي إليه . وتشير النتائج السابقة أيضاً إلى أن الإناث داخـــل تلـــك المجموعـــة من الأفراد تميل إلى الانتماء للحزب الوطني بينما يميل الذكور للحزب المعارض . وهذا الاتجاه يعطي دلالة ارتباطية طردية .

معامل التوافق . C

من خلال مناقشاتنا لمعامل الارتباط (فاي) قلنا إن هناك بعض القيود التي تحد مـــن استخدام هذا المعامل . فإن كان معاملي التوافق وفاى متشابهين إلى حد كبير، مـــن حيـــث أهميـــتهما في قياس العلاقة بين متغيرين، يتم قياسهما بواسطة مقياس اســـمي، فـــإن معامل فاي يصلح للبيانات الممكن تصنيفها في جدول مزدوج أو ما

يســـمى بجـــدول (٢ × ٢) ، أما ما يزيد على ذلك فلا يقدر معامل فاي على قياسه ومـــن ثم استطاع كارل بيرسون أن يصور معاملاً آخر أسماه معامل التوافق يمكن استخدامه في جداول أكبر قي تقسيماتها التصنيفية عن (٢ × ٢) مثـــال ذلك جداول (٣ × ٤) أو (٢ × ٥) أن يتضـــمن ثـــلاث تصنيفات للمتغير الأول وأيضًا خمس تصنيفات للمتغير الثاني.

مثال :

يتضـــمن الجدول الآتى بيانات العلاقة بين المستوى التعليمي وعوامل التغيب عـــن العمـــل لمائة عامل في أحد المصانع بشبرا الخيمة . وأراد باحث أن يحسب العلاقة الارتباطية بين المتغيرين.

العلاقة بين المستوى التعليمي وأسباب التغيب عن العمل

الحالة التعليمية	مرض	مشكلات أسرية	مشكلات عمل	المجموع
تعليم جامعى	٥	١٠	٣٥	٥٠
تعليم متوسط	٦	٩	١٥	٣٠
أمى	٨	٧	٥	٢٠
المجموع	١٩	٢٦	٥٥	١٠٠

خطوات حل المثال:

١- تربيع كل تكرار من التكرارات المدونة في خلايا الجدول.

٢- نقســم مــربع كل تكرار حصلنا عليه على حاصل ضرب مجموع تكرارات الصـــف في مجموع تكرارات العمود الواقعة في خانة هذا التكرار وذلك على النحو التالي:

$$\frac{\text{مربع تكرار الخلية}}{\text{مجموع تكرار العمود} \times \text{مجموع تكرار الصف}}$$

٣- جمع خوارج القسمة للحصول على (مجـــ).

٤- حساب معامل التوافق (C) باستخدام المعادلة الآتية :

$$C = \sqrt{١ - \frac{١}{\text{مجـــ}}}$$

حل المثال :

مجـــ الصف الأول $= \frac{(٥)^{٢}}{١٩ \times ٥٠} + \frac{(١٠)^{٢}}{٢٦ \times ٥٠} + \frac{(٣٥)^{٢}}{٥٥ \times ٥}$

$= \frac{٢٥}{٩٥٠} + \frac{١٠٠}{١٣٠٠} + \frac{١٢٢٥}{٢٧٥٠}$

$= ٠,٠٢٦ + ٠,٠٧٧ + ٠,٤٤٥ = ٠,٥٤٨$

مجـــ الصف الثانى $= \frac{(٦)^{٢}}{١٩ \times ٣٠} + \frac{(٩)^{٢}}{٢٦ \times ٣٠} + \frac{(١٥)^{٢}}{٥٥ \times ٣٠}$

$= \frac{٣٦}{٥٧٠} + \frac{٨١}{٧٨٠} + \frac{٢٢٥}{١٦٥٠}$

$= ٠,٦٣ + ٠,١٠٤ + ٠,١٣٦ = ٠,٨٧$

مجـــ الصف الثانى $= \frac{(٨)^{٢}}{١٩ \times ٢٠} + \frac{(٧)^{٢}}{٢٦ \times ٢٠} + \frac{(٥)^{٢}}{٥٥ \times ٢٠}$

$= \frac{٦٤}{٣٨٠} + \frac{٤٩}{٥٢٠} + \frac{٢٥}{١١٠٠}$

$= ٠,١٦٨ + ٠,٠٩٤ + ٠,٠٢٣ = ٠,٢٨٥$

مجموع الصفوف $= ٠,٥٤٨ + ٠,٨٧ + ٠,٢٨٥ = ١,٧٠٣$

وباستخدام المعادلة السابقة يكون معامل التوافق كالآتي:

$$C = \sqrt{١ - \frac{١}{١,٧٠١}} = ٠,٦٤٣$$

ملاحظات عامة على معامل التوافق:

١- لا يصـــلح معامل التوافق للمقارنة بين جدولين إلا بشرط واحد هو أن تتساوى أعداد الصفوف والأعمدة بينهما .

٢- يســتخدم لقــياس ارتــباط بين متغيرات كيفية ويمكن استخدام معاملى ارتباط التوافق وفاى فى حالة المتغيرات التى تنقسم إلى فئات كمية.

المفاهيم الأساسية Key Concepts

١- ارتباط خطي بسيط :

يقاس بمعامل الارتباط وقيمته تتراوح بين (+١) ، (-١) . وأساس حسابه هو استـــخدام انحـــرافات كل من المتغيرين التابع والمستقل عن المتوسط الحسابي لكل منهما . وهي طريقة بيرسون في حساب معامل الارتباط .

٢- معامل سبيرمان للرتب :

يعـــد حالـــة خاصـــة من معامل ارتباط بيرسون (ر) إلا أنه يستخدم في حالة العيـــنات صـــغيرة الحجـــم ويقـــيس العلاقـــة الارتباطية بين متغيرين يتم قياسهما بالمقاييس الرتبية Ordinal Scales.

٣- معامل الارتباط فاى :

وهـــو حالـــة خاصـــة مـــن معـــامل بيرسون للارتباط ويستخدم عندما يكون المتغـــيران (س ، ص) ذات بيانات اسمية nominal وتكون بيانات كل منهما من النوع الاسمي الثنائي أي يمكن تصنيفها في جدول مزدوج (٢ × ٢).

٤- معامل التوافق :

يســـتخدم فـــي حالة المتغيرات ذات البيانات الاسمية والتي تصنف في جداول ذات تقسيمات نوعية أكبر من

(٢ × ٢) مثال ذلك جداول (٣ × ٤) أو (٣ × ٥) ..... إلخ.

ملحوظة: تمارين هذا الفصل فى نهاية الفصل الثامن مع تمارين (الانحدار).

٨

الفصل الثامن

الانحدار الخطى

أولاً: أهـــم الطرق الشائعة فى دراسة الانحدار من البيانات الخام

١- الشكل الانتشارى.

٢- طريقة المربعات الصغرى.

ثانيًا: الانحدار المتعدد.

الفصل الثامن
الانحدار الخطى

مقدمة :

يــرجع استخدام لفظ "انحدار" من الناحية الإحصائية إلى عام ١٨٨٥ وذلك عندما استخدمه فرنسيز جالتون Gailton في مقاله الذي نشره خلال ذلك العام، والــذي ضمنه نتائج دراسته عن العلاقة بين أطوال الآباء وأبنائهم وبيّن فيه أن هــناك انحــدارًا لطول الأبناء نحو متوسط أطوال المجتمع الأصلي (موضوع الدراســة)، كمــا خلص إلى نتيجة مهمة حين ذكر أن قيم أطوال الأبناء تنحدر نحو موضع ما يقع ما بين أطوال آبائهم والقيمة المتوسطة (المتوسط) للمجتمع الأصــلي. ولقــد استفاد بهذه النتيجة كارل بيرسون فيما بعد بما أسماه معامل الانحــدار. فــإذا كــان معامل الارتباط – كما يرى بيرسون – يعطي تلخيصًا واضحًا للعلاقة بين متغيرين (س، ص)، فإن معامل الانحدار يعبر عن المتغير المــتوقع أي التنــبؤ في المتغير (ص) (بوصفه متغيرًا تابعًا) كلما تغيرت قيم المتغــير (س) المــناظرة على أساس أنه متغير مستقل. ومن ثم يتحدد الهدف الأساســي لمعــامل الانحــدار في قياس تأثير المتغير (س) على المتغير التابع (ص) ووضــع العلاقــة في شكل معادلة انحنائية أو خطية. وما يهمنا في هذا الجــزء هــو الانحدار الخطي، واستخدام معادلة تفسر هذا النوع من الانحدار، والتــي يمكــن صياغتها، فى معادلة من معادلات الدرجة الأولي على الصورة التالية:

ص = أ + ب س

حيــث أ = مقدارًا ثابتًا يساوي قيمة المتغير (ص) إذا كانت قيمة (س) تساوي صفرًا في المعادلة الموضحة. وتقاس قيمة (أ) على المحور الصادي وذلك في حالة انحــدار (ص) علــى (س)، ب = الميل slope لخط الانحدار على المحور الأفقي والــذي يساوي جبريًا ظل زاوية ميل خط الانحدار على المحور الأفقي. كما يمثل أيضًــا كمــية التغير في قيمة المتغير (ص) المصاحبة لكل وحدة تغير من وحدات تغير المتغير المستقل (س).

ويهـدف الفصـل إلـى أن يعرف الطالب المقصود بالانحدار الخطى وكيفيه حسـابه وأهميـته فـى عملية التنبؤ بقيمة أحد المتغيرين وليكن (ص) فى علاقته بالمتغـير (س) معلـوم القيمة. وأن يتقن كيفيه التعبير عن العلاقة باستخدام الشكل الانتشـارى بين المتغيرين (س، ص). وأن يقارن بين الأشكال المختلفة للعلاقة مع إمكانية تحديد اتجاه هذه العلاقة باستخدام الرسم البيانى.

أولاً: أهم الطرق الشائعة في دراسة الانحدار للبيانات الخام :

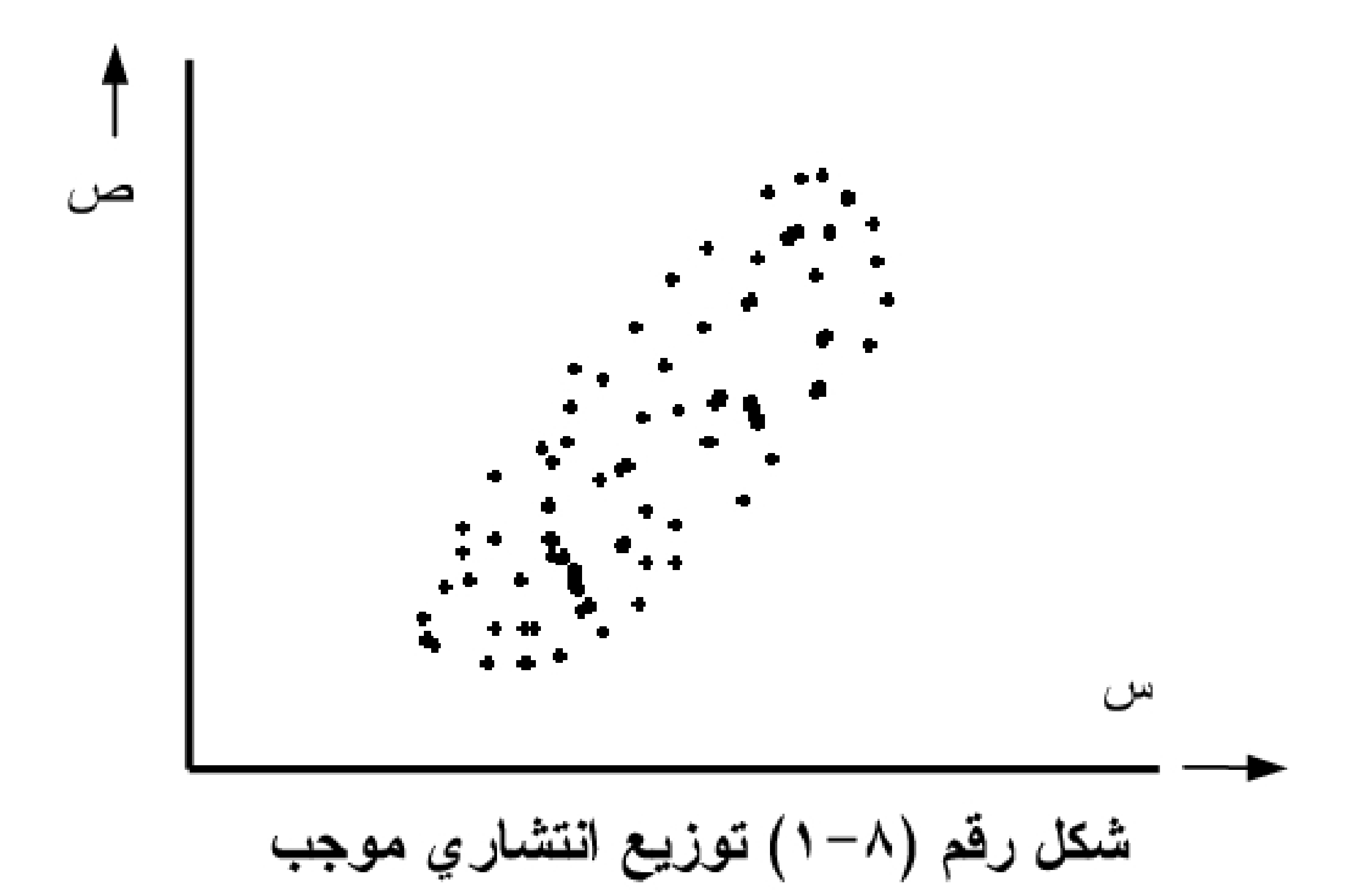
١- الشكل الانتشاري :

ويستخدم هذا الشكل للتعرف مبدئيًا على شكل العلاقة بين المتغيرين (س ، ص) وذلـك باسـتخدام محاور الإحداثيات (المحور السيني والمحور الصادي) حيـث يـتم رصد وتمثيل كل قيم المتغير الأول مع ما يناظرها من قيم المتغير الآخـر فـي نقطة توقع على الشكل، حيث يكون لكل نقطة قيمتان (س ، ص) تحـددان موضـعيهما، ونستمر في رصد جميع النقاط حتى نحصل على شكل انتشـارى لجميع قيم س، وص ويمكن من خلال الشكل الانتشاري تحديد ماهية العلاقة وهل توجد أم تنعدم بين المتغيرين؛ فضلاً عن معرفة اتجاه تلك العلاقة في حالة وجودها.

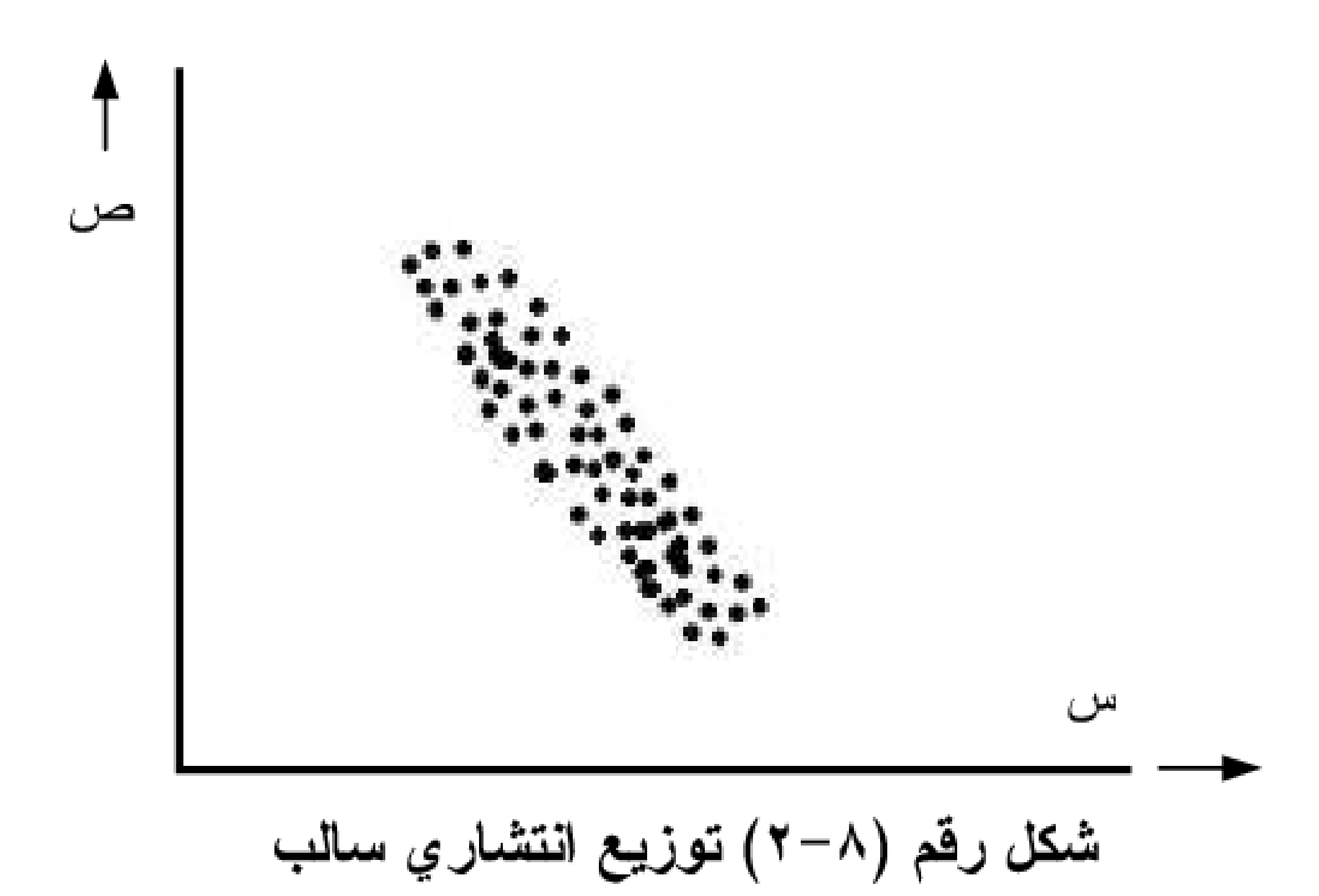
ففـي الشـكل رقم (٨-١) نجد شكلاً انتشاريًا لقيم (س ،ص) يغلب عليها الاتجاه ناحية اليمين، ابتداءًا من ناحية نقطة الأصل. كما يلاحظ وجود نوع من الـتجانس فـي القيم، أي تقل خاصية التشتت. ويطلق على هذا الشكل الانتشار الموجب.

وفـي شـكل رقـم (٨-٢) يتضـح أيضًا وجود تجانس إلى حد ما بين القيم وانخفـاض تشـتتها وتـنافرها، وأن اتجاه الانتشار ناحية اليسار. ويطلق على هذا الشكل الانتشاري السالب.

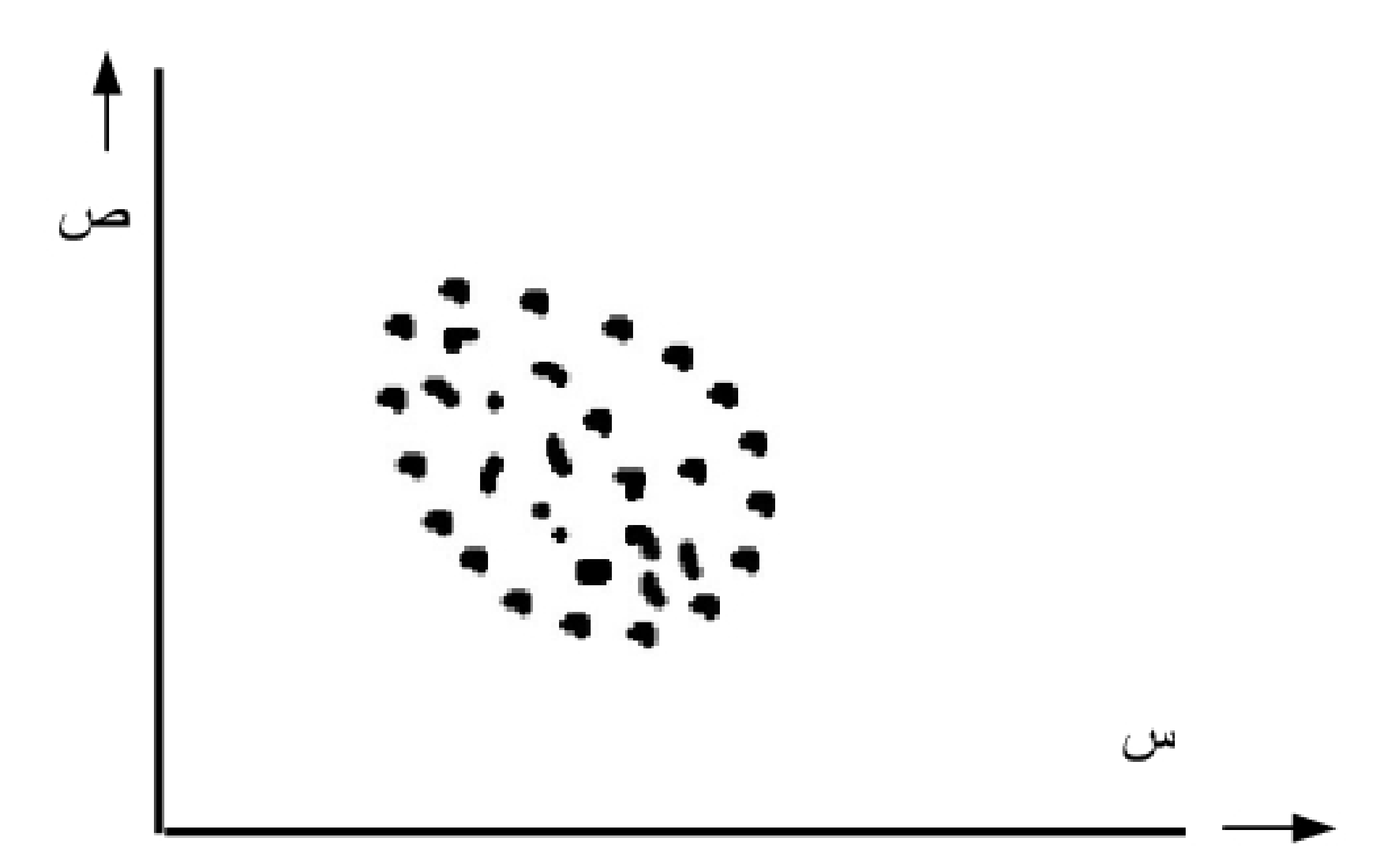
أمـا فـي الشـكل رقم (٨-٣) فنلاحظ عدم انتظام النقط وتشتتها على الشكل الانتشـاري بحيث يصعب رسم خط مستقيم يربط بين معظم تلك النقط، ومن ثم لا توجد أي علاقة بين المتغيرين (س، وص).



شكل رقم (٨-١) توزيع انتشاري موجب



شكل رقم (٨-٢) توزيع انتشاري سالب



شكل رقم (٨-٣) توزيع انتشاري يبين عدم وجود علاقة بين المتغيرين (س ، ص)

مثال :

ارســـم الشـــكل الانتشاري للعلاقة بين المتغيرين (س، ص) من واقع البيانات الموضحة بالجدول الآتى:

جدول رقم (٨-١)

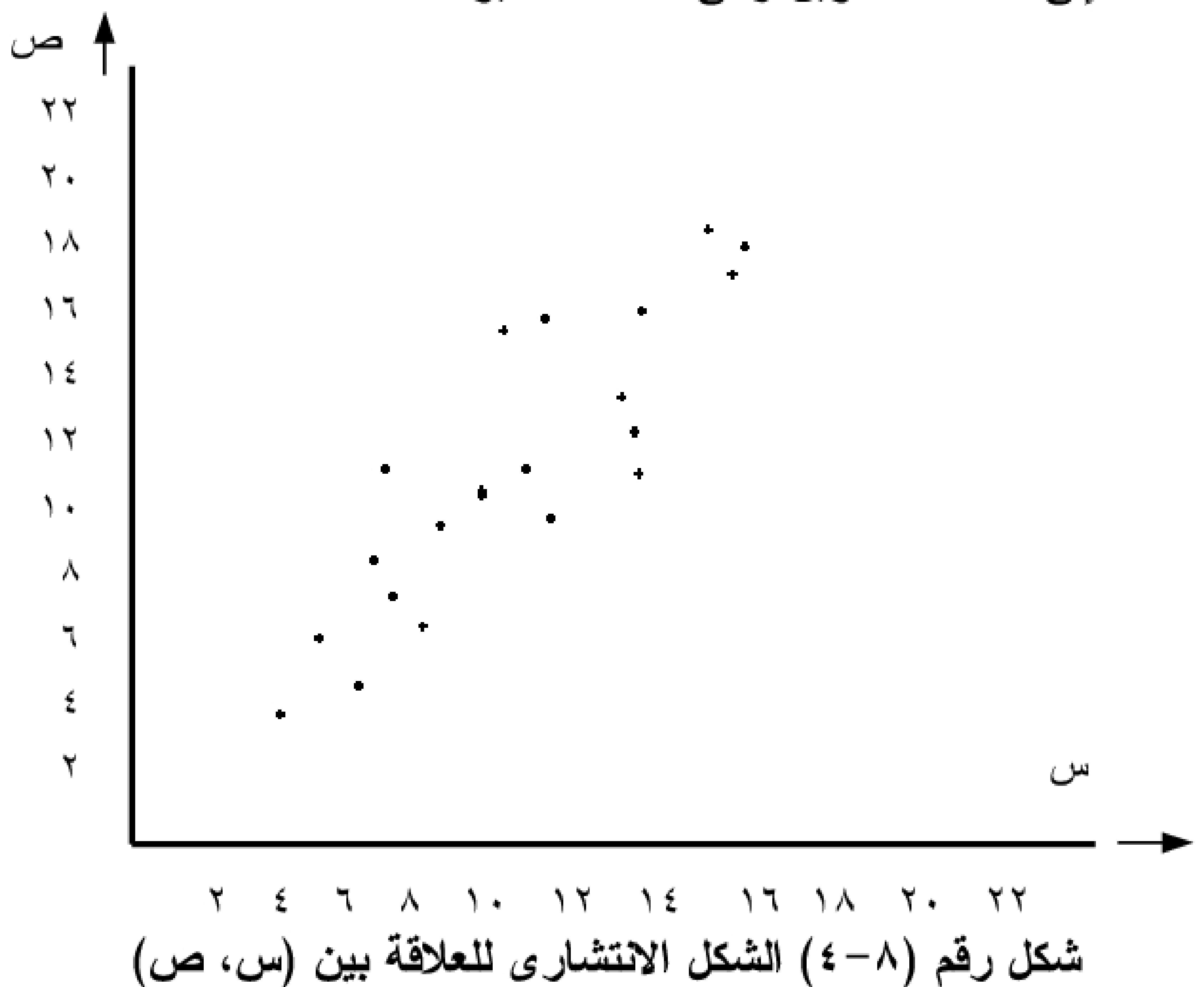
مسلسل	س	ص
١	١٥	١٢
٢	١٠	١٣
٣	٧	٩
٤	١٨	١٨
٥	٥	٧
٦	١٠	٩
٧	٧	١٤
٨	١٧	١٦
٩	١٥	١٠
١٠	٩	١٢
١١	٨	٧
١٢	١٥	١٣
١٣	١١	١٤
١٤	١٧	١٩
١٥	٨	١٠
١٦	١١	١٦
١٧	١٢	١٢
١٨	١٣	١٦
١٩	١٨	١٩
٢٠	٧	١١

الحل :

نرســـم المحورين (س ، ص) ثم نوقع عليها كل قيمة للمتغير (س) وما يناظرها فـــي الجـــدول من قيمة للمتغير (ص)، ونكرر ذلك العمل في العشرين حالة المعطاة فنحصل على عشرين نقطة منتشرة كما هو مبين بالشكل الانتشاري (٨-٤).

مثال:

النقطة الأولى إحداثياتها (س ، ص) هي (١٥ ، ١٢)، فنأخذ بمقياس رسم مناسب قيمة (١٥) على المحور السيني ابتداء من نقطة الأصل، وعند القيمة نقيم خطًا رأسيًا موازيًا للمحور الصادي، وبعد ذلك نأخذ مقياس رسم مناسب على المحور الصادي ونرصد قيمة (١٢) على هذا المحور فتتحدد، ومنها نرسم خطًا أفقيًا في اتجاه المحور السيني وموازي له فيلتقي الخطان الرأسي والأفقي عند نقطة تمثل الحالة الأولى في خانة المسلسل بالجدول. ونكرر العمل بالنسبة لباقي الحالات حتى نصل إلى الحالة العشرين وهى الحالة الأخيرة.



شكل رقم (٨-٤) الشكل الانتشارى للعلاقة بين (س، ص)

ولتحديد خط الانحدار يجب أن نختار خطًا يتوسط جميع النقاط في الشكل الانتشاري السابق، وهناك طريقتان لعمل ذلك، إما أن نقوم برسم هذا الخط بواسطة اليد ولهذه الطريقة عيوبها، حيث يعتمد رسم الخط على مهارة الدارس وإما أن نستخدم طريقة رياضية وهي طريقة المربعات الصغرى.

٢- طريقة المربعات الصغرى :

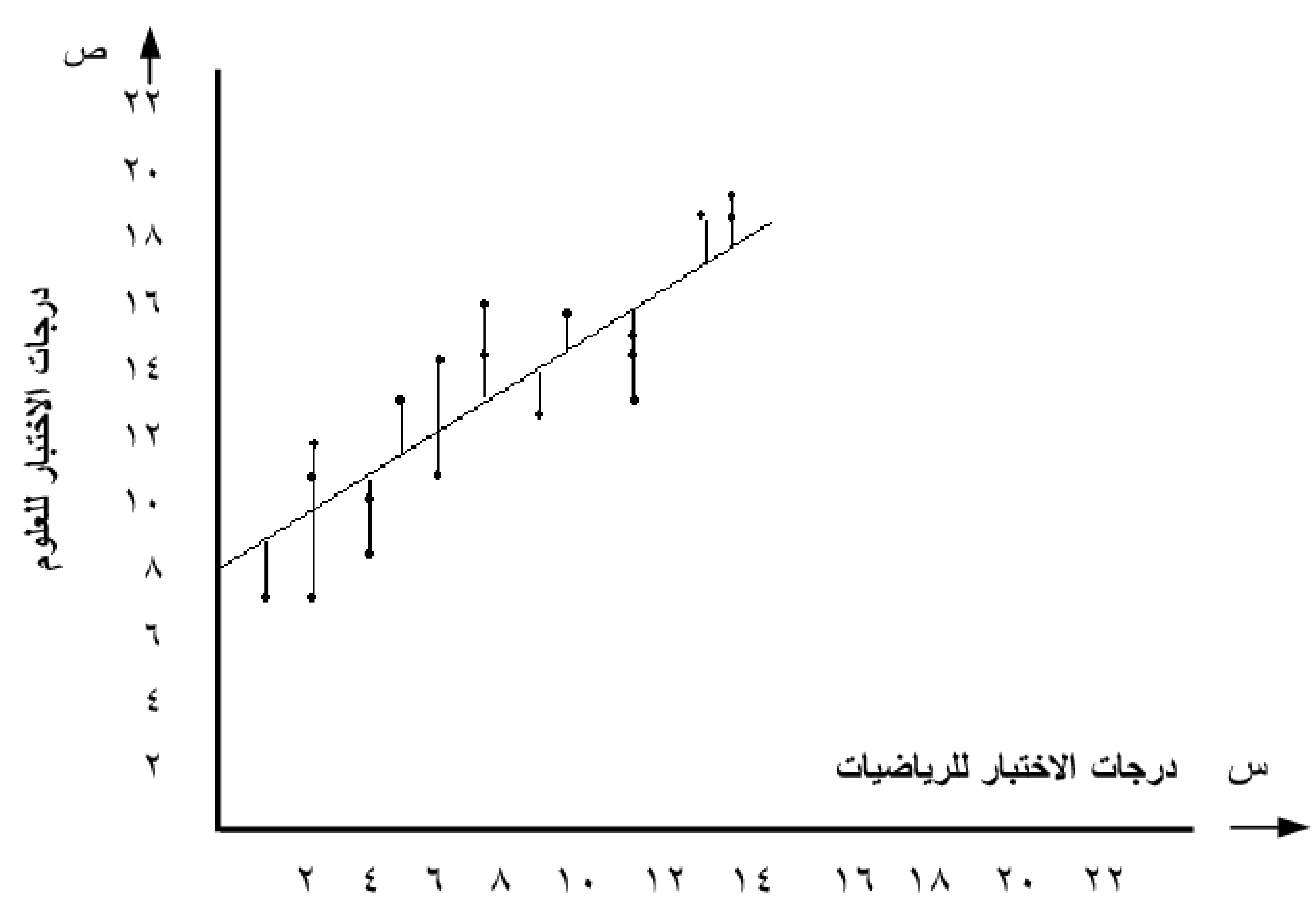
لإبراز العلاقة بين المتغيرين بشكل ملخص ودقيق (س ، ص)، تستخدم طريقة المربعات الصغرى، ويمكن باستخدام تلك الطريقة تمثيل العلاقة بين

المتغيريـــن (س ،ص) بخط مستقيم يمر خلال النقط في الشكل الانتشاري، وأفضل خط مستقيم يمثل الانحدار هو ذلك الذي يمر بمعظم القيم المركزية، أو يمر بالمسار المركزي عبر النقط في الشكل الانتشاري ويعرف المسار المركزي بأنه الخط الذي تكــون قــيمة مجموع مربع المسافات حوله بين النقط أقل ما يمكن. وهذا الخط أو المسار المركزي يعتبر خط الانحدار المنشود.

ومـــن الناحـــية الإحصائية يمكن القول إن خط الانحدار هو خطٌ متوسط يعبر عـــن القـــيم المتناظرة للمتغيرين (س ،ص) بحيث إن مجموع انحرافات قيم (ص) الفعلـــية عـــن قيم المتوسط الحسابي للمتغير (ص) يساوي صفرًا ويمكن أن يلحظ الدارس أن من خصائص المتوسط الحسابي كما ذكرنا سابقًا أن تكون قيمة مجموع مربع انحرافات القيم الفعلية عن المتوسط أقل ما يمكن.

ونخلـــص مما سبق أن خط الانحدار للشكل الانتشاري يلعب دورًا بمثابة نقطة الاتزان للتوزيع الثنائي المرتبط فضلاً عن فائدته فى التنبؤ بقيم المتغير التابع (ص) في علاقته بالمتغير المستقل (س).

وفـــي المثال السابق لو قمنا بتوصيل خطوط رأسية بين النقط على جانبي خط الانحـــدار لوجدناهـــا قريبة جدًا من هذا الخط، وبشكل أقرب للانتظام منه للانتشار والتفرق كما يتضح ذلك من الشكل رقم (٨–٥).



شكل رقم (٨–٥) خط الانحدار باستخدام خاصية المربعات الصغرى.

معادلة انحدار ص على س :

قلنا إن طريقة المربعات الصغرى تعطي أكثر الخطوط توفيقًا لانحدار المتغير الأول وليكن (ص) على المتغير الثاني (س). وإن معادلة هذا الخط تكون على الصورة:

ص = أ + ب س .... معادلة رقم (١)

وتسمى بمعادلة خط الانحدار (ص على س) حيث (أ) هي الجزء المقطوع intercept من المحور الصادي، و(ب) هي ميل خط الانحدار.

مثال :

يوضح الجدول الآتى توزيع الدخل اليومى لعينة مكونة من اثنى عشر عاملاً وأيضًا درجاتهم فى الرضا عن العمل. والمطلوب إيجاد معادلة الانحدار الخطى ثم رسم خط انحدار ص على س أو الرضا عن العمل على الأجر اليومى لهؤلاء العمال الاثنى عشر.

العمال	الأجر اليومي (س)	الرضا عن العمل (ص)
١- محمد	١٠,٥٠	٩١
٢- أحمد	٩,٥٠	٨٩
٣- علية	٩,٠٠	٨٩
٤- حسين	٨,٢٥	٩٠
٥- منال	٨,٠٠	٨٤
٦- زينب	٧,٥٠	٩٢
٧- ماهر	٦,٢٥	٨٦
٨- على	٦,٠٠	٨١
٩- ولاء	٥,٧٥	٨٦
١٠- طارق	٥,٥٠	٨٢
١١- فاطمة	٤,٥٠	٧٤
١٢- حامد	٤,٢٥	٨١

الحل :

المتغير المستقل هو الأجر اليومي (س).

المتغير التابع هو الرضا عن العمل (ص).

مــن المعادلــة ص = أ + ب س نوجــد قيمتــي (أ) ، (ب) من المعادلتين التاليتين:

$$ب = \frac{ن\ مجـ\ س\ ص - (مجـ\ س)\ (مجـ\ ص)}{ن\ مجـ\ س^{٢} - (مجـ\ س)^{٢}}$$ .... معادلة رقم (٢)

$$أ = \frac{مجـ\ ص - ب\ مجـ\ (س)}{ن}$$ .... معادلة رقم (٣)

وللحصــول علــى قيم (مجـ س ص)، (مجـ س)، (مجـ س)، (مجـ $س^{٢}$) **نقوم بعمل الجدول الآتي :**

جدول رقم (٨-٢)

العمال	الأجر (س)	الرضا عن العمل (ص)	$س^{٢}$	س ص
١	١٠,٥٠	٩٤	١١٠,٢٥	٩٨٧,٠٠
٢	٩,٥٠	٨٩	٩٠,٢٥	٨٤٥,٥٠
٣	٩,٠٠	٩١	٨١,٠٠	٨١٩,٠٠
٤	٨,٢٥	٩٠	٦٨,٠٦	٧٤٢,٥٠
٥	٨,٠٠	٨٤	٦٤,٠٠	٦٧٢,٠٠
٦	٧,٥٠	٩٢	٥٦,٢٥	٦٩٠,٠٠
٧	٦,٢٥	٨٦	٣٩,٠٦	٥٣٧,٥٠
٨	٦,٠٠	٨١	٣٦,٠٠	٤٨٦,٠٠
٩	٥,٧٥	٨٦	٣٣,٠٦	٤٩٤,٥٠
١٠	٥,٥٠	٨٢	٣٠,٢٥	٤٥١,٠٠
١١	٤,٥٠	٧٤	٢٠,٢٥	٣٣٣,٠٠
١٢	٤,٢٥	٨١	١٨,٠٦	٣٤٤,٢٥
	٨٥,٠٠	١٠٣٠	٦٤٦,٤٩	٧٤٠٢,٢٥

$$ب = \frac{١٢ (٧٤٠٢,٢٥) - (٨٥) (١٠٣٠)}{١٢ (٦٤٦,٤٩) - (٨٥)^{٢}}$$

$$= \frac{٨٨٨٢٧ - ٨٧٥٥٠}{٧٧٥٧,٨٨ - ٧٢٢٥}$$

$$= \frac{١٢٧٧}{٥٣٢,٨٨} = ٢,٣٩٦$$

$$أ = \frac{١٠٣٠ - (٢,٣٩٦) (٨٥)}{١٢}$$

$$= \frac{١٠٣٠ - ٢٠٣,٦٦}{١٢}$$

$$= \frac{٨٢٦,٣٤}{١٢} = ٦٨,٨٦$$

ويمكـــن اســـتخدام القيمتين (أ) ، (ب) في رسم خط الانحدار بأن نبدأ بإيجاد قيمة (ص) عند (س) = صفر من المعادلة رقم (١).

$\hat{ص}$ = ٦٨,٨٦ + ٢,٣٩٦ (صفر)

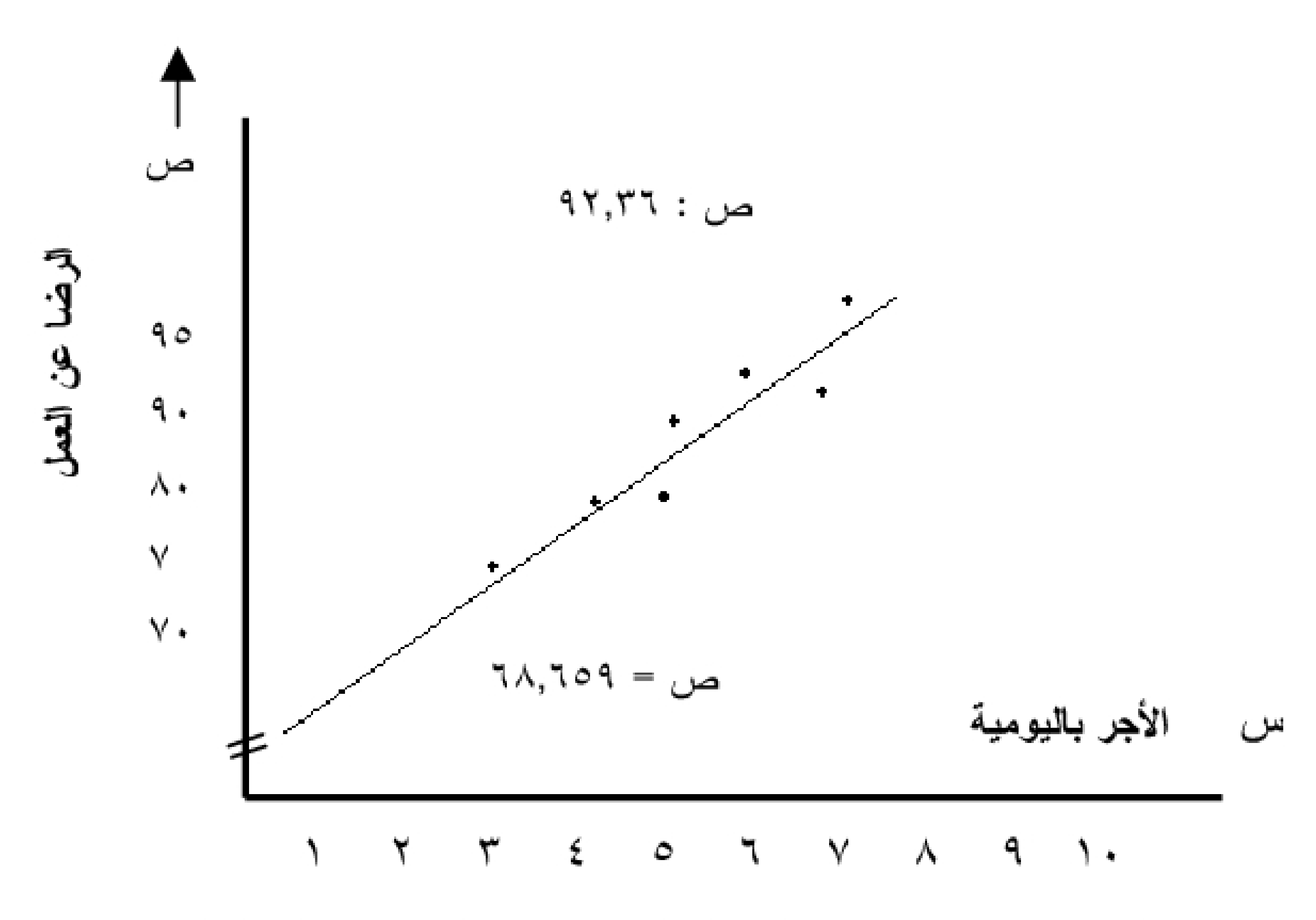
= ٦٨,٨٦

ثم نوجد قيمة ص على س عندما تكون س = ١٠

$\hat{ص}$ = ٦٨,٨٦ + ٢,٣٩٦ x ١٠

= ٩٢,٨٢

وهكـــذا نرسم خط الانحدار مع ملاحظة أن قيمة ص = ٩٢,٨٢ سوف تقع علـــى هـــذا الخـــط. وسوف نحصل على خط الانحدار كما يصوره الشكل رقم (٨-٦).



شكل رقم (٨-٦) خط انحدار الرضا عن العمل على الأجر

ثانياً: استخدام الانحدار المتعدد

ومن الممكن أن يستخدم الباحث الانحدار المتعدد إذا أراد أن يعرف تأثير عدد من المتغيرات المستقلة على المتغير التابع (موضوع الدراسة). ويتم قياس التأثير النسبى لكل متغير مستقل فى المتغير التابع بعد التحكم فى باقى المتغيرات المستقلة الأخرى. ويتم تحديد هذا التأثير من خلال قيم بيتا (β). ويستخدم برنامج (SPSS) فى هذه الحالة.

مثال :

يوضح استخدام الانحدار المتعدد والتعليق على النتائج ففى الجدول الآتى (٨-٣): نتائج دراسة أجريتها فى إحدى المدن الجديدة بدولة قطر عام١٩٩٣[*] وتمت الاستعانة بالانحدار المتعدد للكشف عن تأثير عدد من المتغيرات المستقلة على علاقات الصداقة بين المقيمين بالمدينة داخل نطاق الجيرة القريبة (المتغير التابع)، وتعكس قيم (β) التأثير لكل متغير بالجدول مع التحكم فى باقى المتغيرات المستقلة.

---

* اعتماد محمد علام، النمو الحضرى والمدن الجديدة، فى المجتمع القطري، كلية الإنسانيات والعلوم الاجتماعية، جامعة قطر، الطبعة الأولى ١٩٩٣ ص ص ١٧٩-١٨١.

جدول رقم (٨-٣)

الانحدار المتعدد لعلاقات الصداقة بين الأسر على بعض المتغيرات الديموجرافية والتنظيمية والمجتمعية داخل نطاق الجيرة القريبة

العوامل	معامل بيتا (β)
مجال النشاط الاقتصادى	– ٠,٠٨١
النوع	– ٠,٠٢٣
الحالة التعليمية	– ٠,٠٠٢
الحالة الزواجية	– ٠,٠٠٤
السن	– ٠,٠٦٣
تباين الجنسية	– ٠,١٨٥(*)
الفئة الوظيفية	– ٠,١٦٠(*)
وجود أطفال أقل من ٥ سنوات	٠,٠٠٩
وجود أطفال من سن ٥ – ١٠ سنوات	٠,٠١٤
الابناء من سن ١٠ – ١٥ سنة	٠,٠١٦
الابناء من سن ١٥ سنة فأكثر	٠,١٥٩(*)
طول مدة الاقامة بالمدينة	٠,١٥٦(*)

قيمة ف = ١٥,٥٩ دالة إحصائية عند مستوى (٠,٠٠١) ن = ١٦٥

(*) التأثير المباشر دال إحصائية عند مستوى (٠,٠٥)

تكشف نتائج الدراسة المدونة بالجدول السابق أن تعدد الجنسيات فى المنطقة السكنية بمدينة (أمسيعيد) بدولة (قطر) بما يتضمنه من تباين فى اللغة والعادات والتقاليد يؤثر عكسيًا على تكوين علاقات الصداقة الحميمة على مستوى الجيرة القريبة ، ويأتى هذا المتغير فى المرتبة الأولى من حيث قوة التأثير العكسية على تكوين علاقات الصداقة حيث بلغت قيمة β (-٠,١٨٥) يأتى ذلك من حيث التأثير العكسى والفئة الوظيفية فى تكوين علاقات الصداقة الحميمة داخل نطاق الجيرة القريبة. وكان التأثير المباشر لتباين الجنسية أقوى من تأثير الفئة الوظيفية على تكوين هذا النمط من علاقات الصداقة. فبينما قيمة معامل β للعامل الأول (-٠,١٨٥) بلغت (-٠,١٦٠) للعامل الثانى. بمعنى أنه كلما تباينت جنسيات المقيمين داخل نطاق الجيرة القريبة تقل علاقات الصداقة الحميمة بينهم على مستوى الجيرة المباشرة. كما نجد التأثير العكسى لتباين الفئة الوظيفية على تكوين

علاقات الصداقة الحميمة. أى تكون العلاقات الحميمة أقوى بين المتجاورين داخل المناطق السكنية المخصصة لإسكان الفئة المتوسطة(*). عنها بين المتجاورين داخل المـــناطق السكنية المخصصة لإسكان عائلات كبار الموظفين (حيث تقيم كل أسرة في فيلا مستقلة).

علـــى صــــعيد آخر، يؤثر عاملا الصداقة بين الأبناء في سن ١٥ سنة فأكثر وطـــول مـــدة الإقامة بالمدينة، إيجابيًا على تكوين علاقات الصداقة الحميمة داخل نطـــاق الجيزة القريبة. حيث بلغت قيمة معامل β للعامل الأول (٠,١٥٩)، وللعامل الثاني (٠,١٥٦). كما كان التأثير المباشر لهما على تلك العلاقات له دلالة إحصائية عند مستوى (٠,٠٥).

ومـــن ثم يشير جدول الانحدار المتعدد لعلاقات الصداقة الحميمة بين المقيمين بالمديـــنة علـــى العوامل الديموجرافية، التنظيمية، والمجتمعية داخل نطاق الجيرة القريبة، إلى أن أربعة عوامل هي: تباين الجنسية، والفئة الوظيفية، والأبناء فى سن ١٥ سنة فأكثر، وطول مدة الإقامة كان لتأثيرها دلالة إحصائية على تكوين علاقات الصداقة الحميمة عند مستوى دلالة (٠,٠٥) مع اختلاف إتجاه هذا التأثير. أما باقي العوامـــل الموضحة بالجدول فكان تأثيرها المباشر على علاقات الصداقة الحميمة ضعيفًا وغير دالٍ إحصائيًا.

---

* حيـــث تقـــيم هـــذه الفئة فى عمارات تتألف من عدة طوابق يوجد بكل طابق أربع شقق مما يزيد من درجة التفاعل بين السكان.

تمارين على الارتباط والانحدار

١- فـيما يلـي الدخـل الشهري (بمئات الجنيهات) يمثله المتغير (س) لعينة من الأسرالمصرية، ودرجات التحصيل العلمي لأبنائهم (المتغير ص).

الدخل (س)	١٢	١٤	٧	٩	٨	١١	١٣	١١	٦
التحصيل الدراسي للأبناء (ص)	١١	١٥	٥	٨	٨	٩	١٢	١٣	٥

المطلوب:

(١) حساب معامل بيرسون للارتباط.

(٢) تنـبأ بدرجة أحد الأبناء في التحصيل الدراسي إذا كان دخل أسرته (س) في الشهر ١٨٠٠ جنيهًا.

٢- فـيما يلي تقديرات عينة من الطلبة في امتحان مادتي الإحصاء والرياضيات والمطلوب حساب معامل سبيرمان بين تقديرات المادتين.

رقم الطالب	١	٢	٣	٤	٥	٦
الإحصاء	ض	م	جـ	ض جـ	ل	جـ جـ
الرياضيات	ل	جـ جـ	جـ	ض	ض	م

٣- أمكن التوصل إلى البيانات الآتية عن المتغيرين (س) ، (ص)

مجـ س = ٤٨   مجـ ص = ٧٦

مجـ س^{٢} = ٥٣٦   مجـ ص^{٢} = ١١٠٨

مجـ س ص = ٧٢٠   ن = ٦

والمطلوب حساب معامل بيرسون للارتباط بين قيم س ، ص.

٤- باستخدام المعادلة العامة لخط الانحدار (ص = أ + ب س).

احسب قيمة (ص) من البيانات التالية :

(أ) س = ١٢ ، أ = ٥ ، ب = ٣

(ب) س = ٣,٥ ، أ = ٢,٧ ، ب = ٨,١

(ج) س = ١٤ ، أ = ٩ ، ب = ٤

(د) س = ١٠ ، أ = ٤,٥ ، ب = -٣

(هـ) س = ١١,٢ ، أ = ٣ ، ب = ٥,٥

٥- لدراســة العلاقــة بين الدخل (ص) بمئات الجنيهات والاستهلاك (س) بمئات الجنــيهات فــى احدى المناطق السكنية بمدينة القاهرة ، أخذت عينة من(٤٠) أسرة فأعطت النتائج الاتية:

مجـ س = ١٠٠   مجـ ص = ١٢٠

مجـ س ص = ٥١٦   مجـ $س^٢$ = ٤١٠

مجـ $ص^٢$ = ٧٢٠

والمطلوب حساب:

(أ) معامل الارتباط بين الدخل والاستهلاك.

(ب) خط انحدار الدخل على الاستهلاك.

(ج) قيمة الدخل عندما يبلغ الاستهلاك ٧٠٠ جنيهًا.

المراجـــع

المراجع

أولاً: المراجع العربية :

١- أحمـــد عبادة سرحان: مقدمة فى الإحصاء الاجتماعى، الجزء الأول - جامعة الإسكندرية، كلية التجارة.

٢- اعـــتماد علام ويسرى، أساسيات الإحصاء الاجتماعى، دار قطرى بن الفجاءة للنشر والتوزيع، الدوحة، قطر، ١٩٩١.

٣- السيد محمد خيرى: الإحصاء فى العلوم النفسية والتربوية والاجتماعية، الطبعة الثالثة، القاهرة، مطبعة دار التأليف، ١٩٦٤.

٤- زكـــريا الشـــربينى، الإحصاء اللابارامترى، مكتبة الأنجلو المصرية، القاهرة، ١٩٩٠.

٥- صـــلاح أحمـــد مـــراد، الأســـاليب الإحصـــائية فى العلوم النفسية والتربوية والاجتماعية، القاهرة، مكتبة الأنجلو المصرية، ٢٠٠٠.

٦- عـــبد اللطيف عبد الفتاح وأحمد محمد عمر: مقدمة الطرق الإحصائية، الطبعة الـــرابعة، القاهـــرة، شركة الطوبجى للطباعة والنشر، ١٩٨١، ١٩٨٢.

٧- عـــبد الرحمن بن محمد سليمان أبو عمه، وأنور أحمد محمد عبد الله، ومحمود هندى، الإحصاء التطبيقى، ١٩٩٠.

٨- عـــبد الله عـــبد الحلـــيم أبـــو بكـــر، وداود سليمان المدنى، وإسماعيل سليمان العوامرى: أساليب البحث الإحصائى، القاهرة، التجارة والتعاون للطبع والنشر، ١٩٨٤.

٩- فاروق عبد العظيم وآخرون: مبادئ الإحصاء الوصفى والتحليل، الإسكندرية، دار المطبوعات الجامعية، ١٩٨٤.

١٠- فـــاروق عـــبد العظيم، وبدر الدين المصرى: الإحصاء، القاهرة، دار الكتب الجامعية.

١١- فـــتحى عـــبد العزيز أبو راضى، مبادئ الإحصاء الاجتماعى، الجزء الأول، الإسكندرية، دار المعرفة الجامعية، (بدون تاريخ).

١٢- محمـــد سمير إبراهيم وأبو بكر أحمد حسين: أساسيات علم الإحصاء، الجزء الأول، الطبعة الثانية، القاهرة، مكتبة عين شمس، ١٩٦٧.

١٣- ممـــدوح الصـــدفى محمـــد، ومحمد عبد السميع عثمان، وإكرام سيد غلاب، مقدمـــة فـــى الإحصاء الاجتماعى والدراسات الوصفية (الناشر غير مبين).

١٤- محمـــود السـيد أبو النيل: الإحصاء النفسى والاجتماعى والتربوى، المؤسسة الإبراهيمية، ١٩٩٨.

١٥- مصـــطفى رزق: الكمبيوتر للمبتدئين، الطبعة الثانية، أسيوط مكتبة الطليعة، ١٩٨٦.

ثانياً: المراجع الأجنبية :

16- Anderson. T.W. and Stanley L - Sclove, Statistical Analysis of Data. Boston: Houghton Miffin Company, 1978.

17- Blalock. Hubert M, Social Statistics, 2nd. Edition. New York: Mc Graw-Hill Book Company, 1972.

18- Bogue, Donald J. Principles of Demography. New York: John Wiley and Sons. Inc. 1969.

19- Ferman S. Gerlad and Jack Levin, Social Science Research: A Handbook for Students. New York: John Wiley and Sons, 1975.

20- Felice, L., Statistics: a Tool for Social Research, Blemont, California: Wadswarth Publishing.

21- Graham, Alan, Statistics, London: Hodder Headline Plc., 1994.

22- Hinkle, Dennis, Wiersm, William and Jurs, G. Stephen, Applied Statistics for the Behavioral Sciences. Chicago: Rand Mc Nally College Publishing Company, 1979.

23- Kurtz, Norman R. Introduction to Social Statistics Tokyo: Mc Graw-Hill Book Company, 1983.

24- Lutz, Gene M. Understanding Social Statistics. New York: Macmillan Publishing Co. Inc. 1983.

25- Nie, H. Normam et al., Statistical Package for the Social Sciences. New York: Mc Graw-Hill Book Company, 1978.

26- Ott, Lyman. An Introduction to Statistical Methods and Data Analysis - North Scituate. Massachusetts: Duxbury Press, 1977.

27- Parsons, Robert. Statistical Analysis: A Decision - Making Approach. New York: Harper and Raw Publishers, 1979.

28- Robert Niles. Com, Statistics / http://www.robertniles.com/ stat/medianshtml26/02/2006.

29- Shryock. Henry S. and Jacob S. Siegel, The Methods and Materials of Demography. U.S. Department of Commerce. Vol. 1, 1980.

30- Startup: Richard and Ewyn T. Vhittaker. Introducing Social Statistics. London: George Allen and Unwin, 1982.

فهرس الكتاب

الفصل الرابع
مقاييس النزعة المركزية

الفصل الخامس
مقاييس التشتت

الفصل السادس
المنحنى الاعتدالى والمعايير والالتواء

الفصل السابع
الارتباط

الفصل الثامن
الانحدار الخطى